Journal of
Experimental Psychology

VOL. 32, NO. 5 MAY, 1943

RETROACTIVE FACILITATION AS A FUNCTION OF DEGREE OF GENERALIZATION BETWEEN TASKS [1]

BY R. JANE HAMILTON

Smith College

INTRODUCTION

Although retroaction experiments have generally shown only a negative or inhibiting effect from interpolation of a second task upon recall of a first task, it has nevertheless been assumed that under some conditions a positive or facilitating effect may occur. The present experiment was designed to determine whether such facilitation of retention may be obtained and some of the conditions under which it may occur.

The assumption of facilitating effects has been made in the Skaggs-Robinson version of the expected relationship between retroaction and degree of similarity between original and interpolated tasks and more recently in Gibson's application of generalization and differentiation to verbal learning (3). According to the Skaggs-Robinson function (14) retroactive facilitation should be obtained when there is a very high degree of similarity between the original and interpolated tasks. However, when experimenters have varied degree of similarity, they have not obtained degrees of facilitation but have commonly found only inhibition (1, 8). Attention may therefore be turned to Gibson's proposition.

Speaking in terms of paired associates learning, Gibson (3) has stated that retroactive facilitation may be expected when a high degree of generalization exists between the stimulus items of the original and interpolated activities and the response items of the two activities are identical; and that as the degree of generalization decreases, the degree of facilitation will also decrease. By general-

[1] This research was done at Smith College in partial fulfillment of the requirements for the degree of Master of Arts. The writer wishes to express her thanks to Professor Eleanor J. Gibson for her aid and guidance.

ization is meant the tendency for one stimulus item to be confused with another so that the first evokes the response of the second instead of its own response. That is, if there are two tasks, with stimulus *a* and its response *b* in the first and stimulus *c* and its response *d* in the second, stimulus *a* may evoke response *d* instead of *b* during recall of the first task after interpolation of the second. Confusion of the items during recall has been shown to exist in a situation in which the stimulus items of the two lists were of varying degrees of generalization and the response items were different (4). The tendency during recall of the original list for the stimulus items to evoke response items belonging in the interpolated list increased as the degree of generalization between stimuli in the two lists increased; that is, there was retroactive inhibition, and the degree of retroactive inhibition varied directly with the degree of generalization. This relationship existed because discrimination between the stimulus items of the two lists was necessary for recall of the correct response items, and discrimination became more difficult as the degree of generalization between stimulus items increased.

As yet no test has been made of the effect of varying degree of generalization between stimulus items of the two lists when the response items are identical. In such a situation, according to theoretical expectations, there should be an increasing degree of retroactive facilitation as the degree of generalization increases. Since no discrimination between the stimulus items of the two lists would be required, learning the interpolated list should function as further study of the original list and should do so to an extent depending upon the degree of generalization between the stimulus items of the two lists. Thus if stimulus a_1 in an interpolated list has a high degree of generalization with stimulus *a* in the original list, and a_3 in another interpolated list a low degree, then learning $a_1 \rightarrow b$ should facilitate recall of $a \rightarrow b$ more than should learning $a_3 \rightarrow b$. This proposition has been stated by Gibson as follows: "If generalization occurs between two lists, and if the generalizing items have the same responses, recall or relearning of list 1 will be facilitated as the degree of generalization between the two lists increases." (3, p. 210) The task of the present experiment was verification of this proposition and a comparison of the obtained results with the results from Gibson's experiment (4) in which degree of generalization between the two lists was varied while the responses of the generalizing items were different.

Method

Plan of the Experiment.—This experiment was planned to test the hypothesis that with increasing degree of generalization between stimulus items in original and interpolated tasks,

there should be increasing retroactive facilitation if the response items in the two tasks are identical. Lists of paired nonsense forms and syllables were used for both the original and interpolated learning. Differences in degree of generalization between stimulus items were obtained by the use of a standard set of forms for the original learning and lists of forms of three degrees of variation on them for the interpolated learning. In addition there was one condition in which the forms of the interpolated list were identical with those in the original list, and a control condition in which the subjects did not learn a second list but read a magazine or conversed with the experimenter. By means of these last two conditions upper and lower limits to possible facilitation were established. In all the conditions the subjects learned the same original list and relearned it after a constant interval. A plan of the forms used as stimulus items in the original and interpolated lists of all five conditions is given in Table I. The response syllables

TABLE I

OUTLINE OF THE CONDITIONS OF THE EXPERIMENT

Condition	Learning Forms of Original List	Learning Forms of Interpolated List	Relearning Forms of Original List
I	standard	standard	standard
II	standard	first degree generalization	standard
III	standard	second degree generalization	standard
IV	standard	zero degree generalization	standard
V	standard	no learning	standard

in all the interpolated lists were *identical* with those in the original or standard list. Furthermore, in the interpolated lists a particular syllable was always paired with a variation on the standard form with which it had originally been paired, never with a variation on another form.

Standardization of Material.—In order that a comparison might be made between the results of this experiment and those of Gibson's study (4), the same forms and syllables were used in this experiment as in the earlier one.[2] The following method had been used in selecting the forms. A set of 12 standard forms and a number of variations on these forms were drawn, and, in addition, a group of forms unlike the original set was made. The forms of the standard set were paired with nonsense syllables, and this list was presented to groups of subjects who had to learn to say the appropriate syllable when its form was shown alone. Twenty-four hours later the subjects were given a recall test. But then the list of forms had been changed so as to include only three of the standard forms, and in place of the other standard ones, forms varying from them to a greater or less degree. The frequency with which the subjects responded to the new forms as though they were the standard ones was taken as an indication of the degree of generalization existing between them and the standard forms. Those forms which were most often responded to as though they were the standard ones were combined into a list of first degree generalization, and those which were rarely, if ever, responded to were made into a list of zero degree generalization. Forms which produced an intermediate number of responses were put into a list of second degree generalization. The forms in these lists were paired with syllables to make the learning material for the retroaction experiment.

Procedure.—All of the learning was done individually by the method of right associates. After the subject had learned the original list, a twenty-minute period was interpolated before she recalled and relearned the original list. In three of the experimental conditions the subjects read *Life* or the *New Yorker* or conversed with the experimenter for four minutes and then learned a second list in which the forms were of first, second, or zero degree generalization with those in the original list (Conditions II, III, and IV, respectively). In Condition I, after the four-minute period the subjects continued the learning of the original list. After the subject had finished learning the second list, she again read or conversed until the end of the twenty-minute period, when she was given a recall test of the original list. In the control condition the subjects read or conversed throughout the interpolated period.

[2] The syllables were ones of low association value taken from Hull's list (7).

The pairs of forms and syllables were shown to the subjects, one pair at a time, by means of a Chicago-type memory apparatus. During the learning periods, trials were alternated so that the subject was shown first the forms paired with syllables (learning trial) and then the forms alone (recall trial), to see whether she could remember the appropriate syllables. The apparatus was set so that each pair was exposed for two seconds during the learning trials and so that each form was exposed for two seconds during the recall trials. From trial to trial, the order of presentation of pairs was changed. Since 6 different orders had been prepared, a given order was used only once in every 6 trials. The learning of both the original and interpolated lists was continued until the subject was able to name 8 syllables correctly in one trial. When the subject continued the learning of the original list during the interpolated period (Condition I), she continued the learning only until she was again able to say 8 syllables correctly. Relearning of the original list was carried to the point of one perfect recall of all 12 syllables in all conditions. As measures of retention of the original list, three criteria were used: 1. number of syllables correctly recalled before any relearning trials had been given; 2. number of relearning trials required to name 8 syllables correctly in one trial; 3. number of relearning trials required to name 12 syllables correctly in one trial.

The following directions were given to the subjects. Before learning the original list they were told:

> "You will be shown a group of forms paired with words. Study these pairs so that when a form is shown alone, you can say the word that goes with it. You will be shown only one pair at a time. Do not try to learn the pairs in any particular order, because the order will be changed every time. The point is to associate a particular word with the form with which it always appears. After every presentation of the forms paired with words, you will be shown just the forms alone in order to see whether you can remember the words that go with them. If you do remember them, pronounce them out loud. This will continue until you have learned eight of the twelve pairs."

When the subjects had recalled eight syllables correctly, they were told:

> "We shall return to this list later on and give enough trials for you to finish learning."

Before the interpolated learning the directions to the subjects in Condition I (further study of the original list) were:

> "Now we shall go back to the list and repeat it until you can again say eight of the twelve words. The procedure will be just the same as that we have just followed."

The directions to the subjects in Conditions II, III, and IV (learning a list with syllables of first, second, or zero degree generalization) were:

> "Now there will be a new list of the same type for you to learn. The procedure will be just the same as that we have just followed, and it will be continued until you have learned eight of the twelve pairs."

At the end of the interpolated period the subjects were told:

> "Now we shall go back to the (first) list and repeat it until you have learned all the pairs. We shall start with a test trial to see how much you remember. Then we shall proceed with alternate learning and test trials as we did before."

Subjects.—The subjects were 105 students in elementary psychology classes. They were divided into 5 groups, with 15 in Condition I, 25 in Condition II, 20 in Condition III, 25 in Condition IV, and 20 in Condition V. More subjects were put into Conditions II and IV than into any others because these conditions, which had interpolated lists with syllables of first and zero degree generalization, were deemed the most critical for a test of the hypothesis. The smallest number of subjects was assigned to Condition I, with more learning of the original list, since this condition was important only for setting an upper limit to the amount of facilitation which might be expected from interpolation of a second list.

Results

Trials to Learn the Original List.—The mean number of trials required to learn the original list to a criterion of 8 syllables correct

was 6.73 in Condition I, 7.40 in Condition II, 7.40 in Condition III, 7.44 in Condition IV, and 7.35 in Condition V. The largest difference between means (Conditions I and IV) is only seven-tenths of one trial. Since the *t* statistic for this difference is 0.768, the difference is totally unreliable, and the original learning rates may be assumed not to have influenced retention under the various conditions. Furthermore, it follows that original learning rates in this experiment and in Gibson's experiment (4) are not significantly different, since her rates varied from 6.86 to 7.14 and therefore fall within the range of rates in this experiment.

Trials to Learn the Interpolated Lists.—From Table II it may be seen that the mean number of trials taken to learn the interpolated

TABLE II

TRIALS TO LEARN THE INTERPOLATED LIST TO A CRITERION OF EIGHT SYLLABLES CORRECT IN ONE TRIAL

Condition	I	II	III	IV
Mean trials	1.27	2.08	2.35	3.48
σ_M	0.15	0.24	0.23	0.41

lists to a criterion of eight syllables correct in one trial shows a constant increase from Condition I through Condition IV. In Condition II, in which the forms were of first degree generalization with the original forms, only 2.08 trials were required, but in Condition IV, in which the forms were of zero degree generalization, 3.48 trials were necessary. The difference between the means of Conditions II and IV is reliable at the one per cent level of significance (*t* is 2.935). The finding that the number of trials to learn the interpolated list increased from Condition I through Condition IV is what one would expect according to the theory, for the greater the degree of generalization, the more the second list should function as further learning of the original list. If it functions *as* the original list, fewer trials should be required to learn it than if it does not function in this way.

Retention of the Original List.—Retention scores as measured by the first recall trial are given in Table III, both in terms of mean syllables recalled and the percent recalled in each condition. According to these scores there was a constant decrease in retention from Condition I (continued study of the original list) to Condition IV (zero degree generalization between the stimulus items of the two lists). The *t* statistic of the difference between mean recall scores for Conditions II and IV (first and zero degree generalization) is 2.024 and is reliable at the 5 percent level of significance. All the

work conditions except Condition IV, with stimuli of zero degree generalization, showed *better* retention than the control condition, *i.e.*, showed retroactive facilitation, and the degree of retroaction decreased as the degree of generalization decreased. There was retroactive inhibition only in Condition IV with zero degree generalization, and then there was only a 5 percent inhibition effect.

In comparing the number of trials necessary to reach criteria of 8 and of 12 syllables correct in one trial, medians are given instead of means,[3] since by application of Shen's formula (**5**, p. 52), it was found that the medians are the more reliable measure. In Table II

TABLE III

MEASURES OF RETENTION OF ORIGINAL LIST

Condition	I	II	III	IV	V
Syllables Retained, Recall 1					
Mean	7.93	7.52	7.30	6.52	6.85
σ_M	0.30	0.31	0.37	0.38	0.39
Percent Retained	99	94	91	82	86
Trials to Relearn 8 Syllables					
Median	0.0	0.0	0.5	1.0	1.0
σ_{mdn}	0.19	0.17	0.25	0.23	0.37
Percent Saved	100	100	94	86	83
Trials to Learn 12 Syllables					
Median	4.0	5.0	3.5	4.0	4.5
σ_{mdn}	0.73	0.68	0.45	0.66	0.81

is shown the median number of trials necessary to relearn 8 syllables and to learn 12, and also the percent saved in relearning 8 syllables. The median number of trials necessary for recall of 8 syllables shows a gradient with fewest in Conditions I and II and most in Conditions IV and V. Among the work conditions the percent saved in relearning 8 syllables decreased from 100 percent in Conditions I and II, through 94 percent in Condition III, to 86 percent in Condition IV. Condition V with no interpolated learning showed a saving of 83 percent, a value slightly lower than that which occurred when a second list of zero degree generalization (Condition IV) was learned. The *t* value of the difference between trials to 8 correct in Conditions II and IV is 3.546, which is very significant.

[3] The formula used in finding the standard error of a median was $\sigma_{mdn} = \frac{i\sqrt{N}}{2f_p}$. See Guilford (**5**, p. 52).

Because of the gradient of decreasing retention with decreasing degree of generalization between stimuli of original and interpolated lists, and because of the very significant *t* value for the difference between trials in Conditions II and IV, the hypothesis is strongly supported. Since such conclusive support is found for the hypothesis in terms of trials to relearn 8 syllables, it is probably not important that the trials to learn 12 syllables show no trend. Even the difference of 1.5 trials between Conditions II and III is not significant, since the *t* value of the difference is only 1.840. The fluctuations in trials seem to indicate that by the time the entire list is learned the facilitating effect of the interpolated list has been lost.

Percents of retroactive facilitation and inhibition were calculated according to the mean number of syllables retained in recall 1 and the savings scores for relearning 8 syllables. When recall 1 was used as a basis, the percents were all small. Even in Condition I (continuation of original learning) there was only a 16 percent effect, and in Conditions II and III there were 10 percent and 7 percent effects. Instead of facilitation, Condition IV showed a 5 percent inhibition effect. When savings scores to reach a criterion of 8 syllables correct were used as a basis, the percents of facilitation were a little larger and there was no inhibition. Equal facilitation (20 percent) was found in Conditions I and II, in which the interpolated list was either identical with the original or of first degree generalization. A smaller amount of facilitation (13 percent) was found when the interpolated list was of second degree generalization (Condition III), and a slight amount of facilitation (4 percent) occurred even in Condition IV, in which the interpolated list was of zero degree generalization.

Comparison with Gibson's Results.—While the results of this study showed retroactive facilitation, the results of Gibson's study (4) showed retroactive inhibition.[4] In her experiment retention as measured by the first recall trial increased steadily from 2.93 syllables in Condition I, in which the forms were identical with those in the original list and the syllables were different, to 6.21 syllables in Condition IV, in which the forms were of zero degree generalization and the syllables were different. Results of the two experiments seem to indicate that when the stimulus items are of zero degree generalization, it makes little difference whether the response items

[4] The only difference between the two experiments is that in this experiment response syllables in the interpolated lists were identical with those in the original list, and in Gibson's experiment response syllables were different in the original and interpolated lists; the same stimulus forms were used in both experiments. In Gibson's experiment the interpolated conditions were: I. stimuli identical and responses different, II. stimuli of first degree generalization and responses different, III. stimuli of second degree generalization and responses different, IV. stimuli of zero degree generalization and responses different.

are identical or different, for in this experiment, in which the syllables were identical, 6.52 syllables were retained, and in Gibson's experiment, in which the syllables were different, 6.21 syllables

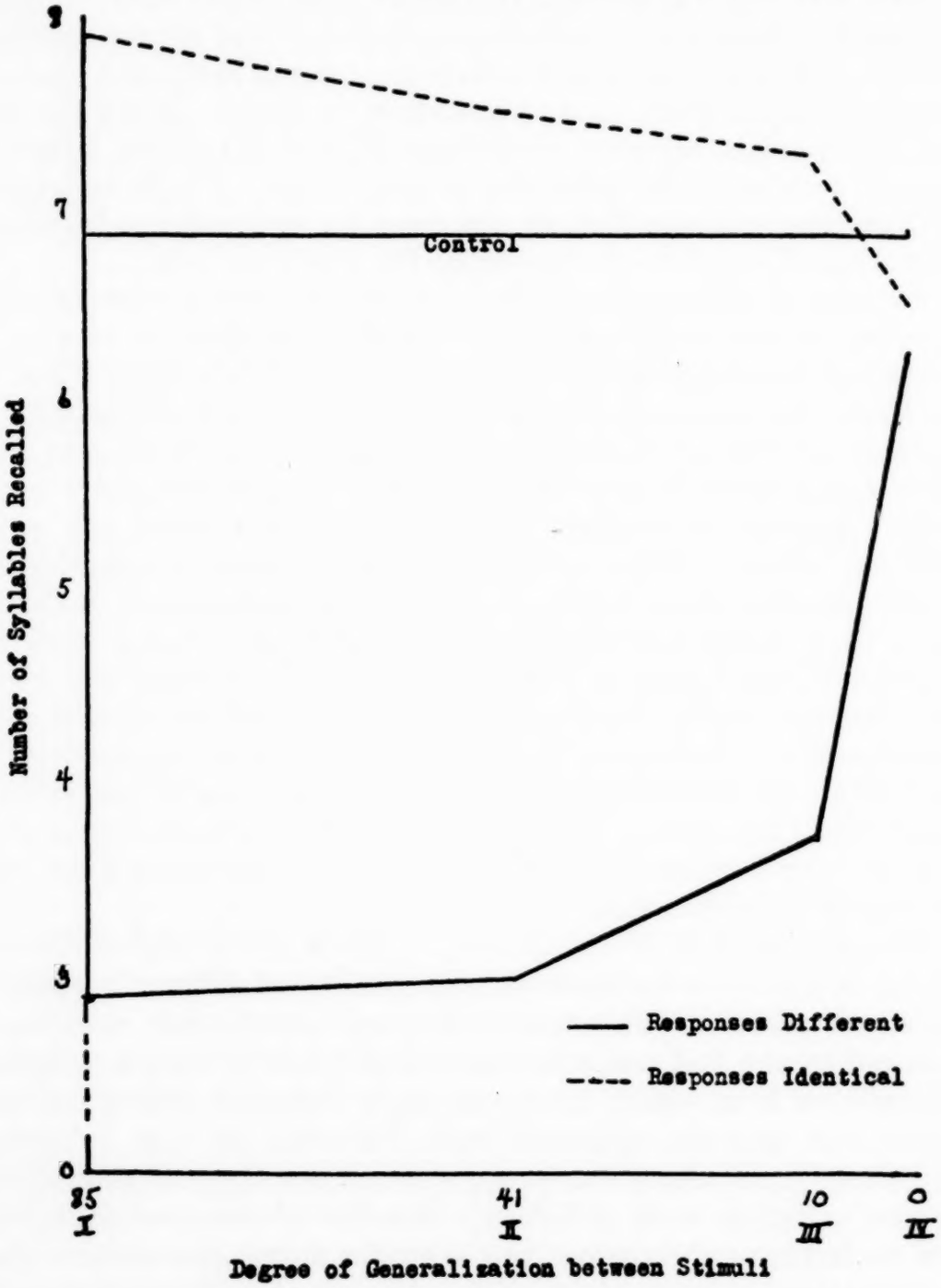

FIG. 1. Retention measured by mean number of syllables in recall 1.

were retained. When response syllables are the same in the original and interpolated lists, the curve of results, in general, shows a gradual decrease in retroactive facilitation with decreasing generalization between the stimulus forms of the two lists (Fig. 1). But when response syllables are different in the two lists, the curve shows

a more rapid decrease in retroactive inhibition with decreasing generalization. The curves of the two experiments meet at the place of zero degree generalization. The gradual slope of the curve of facilitation is not surprising in view of the high degree of retention found in the control condition. Since learning had been carried to a criterion of 8 syllables correct, and 6.85 syllables were recalled in the control condition, facilitation could occur only within the narrow range between 8.00 and 6.85 syllables. From the results of this experiment it is not possible to tell whether the gradual slope is an artifact of the particular experimental conditions or whether it is a characteristic of all curves of facilitation.

In terms of median number of relearning trials required to reach a criterion of 8 syllables correct in one trial, this experiment showed a constant increase in number of trials and the earlier experiment a decrease with decreasing degrees of generalization between stimulus forms.[5] In Gibson's experiment there was a decrease in number of trials from 2.5 in Condition I (forms identical and syllables different) to 1.0 in Condition IV (forms of zero degree generalization and syllables different). In this experiment, on the other hand, there was an increase in trials from 0.0 in Condition I (forms identical and syllables identical) to 1.0 in Condition IV (forms of zero degree generalization and syllables identical). By relearning trials, as well as by syllables in the first recall trial, when stimulus forms were of zero degree generalization, it made no difference whether the response syllables were identical or different in the original and interpolated lists.

Further confirmation of the above-mentioned trends in retention occurs in the number of trials required for learning the interpolated lists in the two experiments. In this experiment the number of trials increased as the degree of generalization between stimulus forms in the two lists decreased. But in Gibson's experiment the number of trials decreased from 7.43 in Condition I (stimuli identical and responses different) to 3.93 in Condition IV (stimuli of zero degree generalization and responses different). Since 3.93 trials were required in her experiment in the condition of greatest ease in learning (Condition IV) and 3.48 trials were required in this experiment in the condition of most difficult learning (Condition IV, with stimuli of zero degree generalization and responses identical), identity of responses in the two lists seems to have been a slight aid in learning in comparison with difference of responses in the two lists. The amount of aid, however, is not great, since there is a difference of only one-half of one trial between the two conditions.

[5] Results as quoted in Gibson's experiment are in terms of means. Medians were calculated in order to make the data comparable with the results of the present experiment.

Discussion of Results

The results of this experiment give conclusive evidence that when response items in a paired associates experiment are identical in original and interpolated lists, retroactive facilitation occurs as a function of degree of generalization between stimulus items of the two lists. For as degree of generalization decreased, so did the amount of *retroactive facilitation* decrease from a maximum in a condition in which the stimulus items of the two lists were identical to a minimum in a condition in which there was a very low degree of generalization between stimulus items. These findings are in contrast to the findings in Gibson's experiment (4) in which the same lists of stimulus items of varying degrees of generalization were used but in which the response items of the original and interpolated lists were different. For in her experiment there were decreasing degrees of *retroactive inhibition* as the degree of generalization between stimulus items decreased. The results of the two experiments together show, therefore, that it is possible to obtain gradients of both facilitation and inhibition by variation in two dimensions—degree of generalization between stimulus items and the necessity or lack of necessity for differentiation between response items of the two lists.

On the other hand, the results of numerous experiments never have shown gradients of both facilitation and inhibition by variation in the single dimension of similarity, as the Skaggs-Robinson function (14) requires. Many of the retroaction experiments, such as Robinson's (14) and Harden's (6) had no control condition and so had no criterion for measuring the presence or absence of facilitation as opposed to inhibition. Other experiments have shown only inhibition (1, 13) or have shown principally inhibition with possibly a very slight degree of facilitation in a single condition of work (9, 15) or facilitation by only one of two measures of retention (8, 12). The retroactive facilitation in these experiments seems so slight as not to be statistically reliable but rather to be due to accident. The only study of variation in similarity in which there may be true facilitation is Dreis's (2); in her study, although there is, in general, an increase in saving scores with an increase in number of identical items in the original and interpolated lists, and although these scores indicate retroactive facilitation in some conditions and inhibition in others, the small range of scores (saving scores 7 percent above and 3 percent below the score in the control condition) together with the lack of a perfect gradient make one hesitate to accept without further evidence the existence of gradients of both facilitation and inhibition by variation in similarity alone.

The effect of variation in two dimensions has been tested by

Watson (16), who found retroactive inhibition in all but two conditions—one in which the original and interpolated tasks were identical and another in which there was a high degree of similarity and perfect 'compatibility' between the two tasks. In these two conditions the facilitation obtained may or may not be reliable, since there is little difference between retention in them and in the control condition.

The principle difference between the Skaggs-Robinson hypothesis and the present hypothesis is that Skaggs and Robinson expected the whole scale of facilitation and inhibition to be produced by variations in degree of similarity alone, while Gibson's theory requires the two variables of degree of generalization between stimulus items of the two lists and necessity or lack of necessity for differentiation between response items. According to her theory, variation in retention may be produced by variation in generalization, but the particular kind of variation is determined by the necessity or lack of necessity for differentiation.

From the number of variables which each hypothesis involves come different methods of graphic representation of findings. In keeping with the implication that one variable will produce both facilitation and inhibition, the Skaggs-Robinson theoretical curve is a single curve. This curve, as is well known, shows first decreasing retention from a point of maximum similarity to some intermediate point in the scale of similarity and then increasing retention to a point of maximum dissimilarity, where, however, retention is not so great as when there is identity of the two tasks. Presumably the first part of the curve of decreasing retention, when similarity is very high, is a curve of retroactive facilitation and the remainder of the curve a curve of retroactive inhibition. On the other hand, the results of the present experiment must be represented in two curves drawn on the same axes, since two factors were varied. And it is necessary to represent the curves of the two conditions as beginning at the same place on the axis showing degrees of generalization, as was done in Fig. 1, for in the two conditions there was a parallel decrease in degree of generalization between the stimulus items of the original and interpolated lists. In these curves there is a continuous decrease in both retroactive facilitation and inhibition; there is no inflection in the curve of inhibition as there is in the Skaggs-Robinson curve. Furthermore, these curves both show a positive acceleration. That is, equal changes in degree of generalization between stimuli produce less change in retention when the degree of generalization is high than when it is low, whether the responses are identical or different in the original and interpolated tasks. On the other hand, the Skaggs-Robinson curve predicts a positive rate of acceleration in the part representing decreasing retroactive facilitation, and first a

positive and then a negative rate of acceleration in the part representing retroactive inhibition. In other words, Robinson apparently expected that when original and interpolated tasks are very similar or very different from each other, equal changes in degree of similarity should produce smaller changes in degree of retention than when the original and interpolated tasks are of an intermediate degree of similarity to each other. This form of inhibition curve has never been found by experiment to agree with obtained results (**2**, **9**, **11**).

Since the Skaggs-Robinson function has never been verified as a whole and Gibson's propositions have been supported by experimental evidence, the conclusion is indicated that a two-variable hypothesis should be accepted in preference to the Skaggs-Robinson function in predicting whether a given set of tasks will show retroactive inhibition or facilitation.

At the same time that the results of the present experiment support Gibson's hypothesis, they also suggest certain problems for further investigation. One of these problems is the relative slopes of the true curves of facilitation and inhibition. Does the true curve of facilitation have a much more gradual slope than the curve of inhibition as the experimental results seem to indicate, or is the difference in slopes only an artifact of the particular experimental conditions used? In order to decide the question some method should be devised by which the percent of retention in the control condition would be reduced below that found in this experiment. For if the difference between the original criterion of learning and retention in the control condition is increased, then there would be a wider range of retention over which facilitation might occur. The best method of producing a greater difference between the original criterion and retention in the control condition might be by an increase in the length of the interval between original learning and recall.

A second problem raised by the findings of this experiment is the role of the response items. For with zero degree generalization between stimulus items of the original and interpolated lists, retention was only slightly better when response items were identical than it was when response items were different. It becomes doubtful whether degree of generalization of response syllables alone, when stimulus forms are kept constant, is important in retention of the original list. An experiment to determine this might test the effect of generalization between response syllables when stimulus forms of the two lists are identical and when the stimuli are of zero degree generalization. In this way it might be seen whether a relationship would be obtained similar to that found when forms are varied and syllables are either identical or different.

Summary and Conclusions

In a retroaction experiment 105 subjects learned lists of pairs of nonsense forms and syllables by the method of right association. Response syllables in the original and interpolated lists were identical in all conditions and stimulus forms in the interpolated lists varied from identity to zero degree generalization with the forms in the original list. Results pointed to the following conclusions:

1. In general, support was found for the hypothesis that in paired associates learning, as the degree of generalization between original and interpolated stimuli increases from zero to a maximum, there will be *increasing retroactive facilitation* if the original and interpolated responses are identical and *increasing retroactive inhibition* if the responses are different. For according to the results, when degree of generalization decreased:

   a. Identity of the response syllables in the two lists resulted in:
      i. decreasing retention of the original list as shown by both recall and relearning,
      ii. decreasing degrees of retroactive facilitation, except for a slight degree of inhibition when the stimulus forms were of zero degree generalization.

   b. A difference in the response syllables in the two lists (as shown by Gibson's experiment (4)) resulted in:
      i. increasing retention of the original list,
      ii. decreasing degrees of retroactive inhibition.

2. When the stimulus forms were of zero degree generalization, there was very little difference in retention in conditions with responses identical and with responses different.

3. Degree of generalization and necessity for differentiation are offered as substitutes for the Skaggs-Robinson hypothesis of decreasing and then increasing retention with decreasing 'similarity.' A graphic representation of the experimental results should be in two curves, since there are two independent variables, instead of in one, like the Skaggs-Robinson curve.

4. The curve for retroactive facilitation was found to have a more gradual slope than the curve for inhibition. Further experimentation is needed to determine whether the difference in slopes of the two curves is characteristic of such curves or whether it is an artifact of the particular experimental conditions.

(Manuscript received December 22, 1942)

References

1. Bunch, M. E., & Winston, M. M. The relationship between the character of the transfer and retroactive inhibition. *Amer. J. Psychol.*, 1936, **48**, 598–608.
2. Dreis, T. A. Two studies in retroaction: I. Influence of partial identity, II. Susceptibility to retroaction at various grade levels. *J. gen. Psychol.*, 1933, **8**, 157–172.
3. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
4. Gibson, E. J. Retroactive inhibition as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.
5. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill Book Co., 1936. Pp. 52, 548–549.
6. Harden, L. M. A quantitative study of the similarity factor in retroactive inhibition. *J. gen. Psychol.*, 1929, **2**, 421–432.
7. Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.
8. Johnson, L. M. Similarity of meaning as a factor in retroactive inhibition. *J. gen. Psychol.*, 1933, **9**, 377–389.
9. Kennelly, T. W. The role of similarity in retroactive inhibition. *Arch. Psychol.*, N. Y., 1941, **37**, No. 260. Pp. 56.
10. Lindquist, E. F. *Statistical analysis in educational research.* Boston: Houghton, Mifflin Co., 1940. Pp. 15–17, 51–54, 56–58.
11. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green and Co., 1942. Pp. 461–469.
12. McGeoch, J. A., & McDonald, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, **43**, 579–588.
13. Matousek, Sister M. A. Reproductive and retroactive inhibition as a function of similarity in the recall and recognition of paired associates. Cath. Univ. Amer., 1939. Pp. 43.
14. Robinson, E. S. The similarity factor in retroaction. *Amer. J. Psychol.*, 1927, **39**, 297–312.
15. Skaggs, E. B. Further studies in retroactive inhibition. *Psychol. Monogr.*, 1925, **34**, No. 161. Pp. 60.
16. Watson, B. The similarity factor in transfer and inhibition. *J. educ. Psychol.*, 1938, **29**, 145–157.

STUDIES OF ABNORMAL BEHAVIOR IN THE RAT

XII. The Pattern of Punishment and its Relation to Abnormal Fixations [1]

BY NORMAN R. F. MAIER AND JAMES B. KLEE

Introduction

When rats are subjected to a discrimination problem on the jumping apparatus, and punishment is applied in a random fashion on 50 percent of their trials (no-solution problem), they soon refuse to perform in the situation. If, however, they are forced to choose, they develop a response which becomes highly sterotyped and persistent. Because responses developed under these conditions have the properties of stereotypy and persistence, and responses developed under conditions of reward do not, it is necessary to designate these two kinds of responses by different terms. Maier, Glaser and Klee (7), and Maier and Klee (8) have called the responses developed in the no-solution situation *abnormal fixations*, to differentiate them from position and discrimination habits. Because an insoluble problem produced a frustrating experience, it became necessary to distinguish between responses developed under conditions of frustration and of motivation.

In one experiment (7), the responses acquired under the two conditions were carefully matched. This was relatively easy to do because most animals formed a response based on position when confronted with the no-solution problem. This position response was then duplicated by purposely training other animals to react to position rather than to the appearance of the stimulus cards.

The positional responses thus matched showed distinct differences when attempts were made to alter them. Although individual differences were present, it soon became apparent that the ordinary type of position response could be broken by so changing the problem that reward accompanied a differential reaction to cards, whereas the other type of position response remained relatively unaffected. Thus, one form of position response could be replaced by training on a discrimination problem, whereas the other could not. It was even shown that animals which persisted in their positional responses ac-

[1] The experiments were aided by grants from The John and Mary R. Markle Foundation of New York City and the Horace H. Rackham School of Graduate Studies of the University of Michigan.

tually learned the discrimination, yet could not overcome the choice in terms of position. The presence of the learned discrimination, which could not come to direct expression, was demonstrated by the manner in which the animal jumped. When the positive card was on the side of the fixated position preference, the rats jumped readily and hit the card squarely, but when the negative card was on the side of the positional tendency, the animals jumped reluctantly and struck the card with the sides of their bodies.

Attempts to break such persistent position responses by (1) punishment for each response, or (2) placing food in an open window on the side opposite, were unsuccessful. In a second experiment (**8**), it was shown that the fixated responses survived a rigorous period of testing during which the animals were subjected to a variety of training situations, conflicts and convulsive experiences.

Only by restraining the animal and physically preventing the appearance of the fixated response (manual guidance) was it possible to produce a breakdown of fixated responses. As soon as the fixated response was removed, the animals, without further training, performed perfectly on the discrimination problems.

This form of positional response showed qualitative differences from ordinary habits in that (a) it had the appearance of a compulsion; (b) it was associated with frustrating conditions; and (c) it was not influenced by the motivating conditions which influence ordinary habits. These characteristics led the authors to designate such behavior by the term, *abnormal fixation*.

The implication of a qualitative distinction in overtly similar responses opens a broad field for further investigation. It, therefore, becomes necessary to explore the fixations in greater detail and determine whether they occur in a great number of learning situations. Kleemeier (**3**) showed that electrical shock at the choice-point of a five-alley discrimination box produced fixations and thus determined an animal's choices. All trial and error learning involves success and failure, and when failure is sufficiently persistent, fixations rather than habits may be formed. These fixations may have remained undetected in the past because they have not been experimentally separated from ordinary habits. It seems desirable, therefore, to investigate the learning of different problems and introduce a variety of patterns of reward and punishment. Thus problems may vary in degree of difficulty, and punishment may be applied 50 percent of the time or 100 percent of the time. If 50 percent punishment is used, it may be applied in a random fashion or it may follow a definite pattern.

Another type of question concerns the influence of one fixation on the next. If a fixation is broken, is the animal more likely or less

likely to form a second fixation? Further, what factors determine the responses which will become fixated? Since frustration and regression have been closely associated, one may expect regression to be related to fixation. If regression is historical in nature, then past experience and native preferences may influence the form the fixation will take. If regression is characterized by de-differentiation or the trend to the more primitive reaction tendencies, then one might expect forces and tensions within the body to give the fixation its specific character. It is quite possible, however, that fixation is unrelated to regression. In this case quite different factors would determine the fixation. The responses existing at the time of frustration or any number of environmental conditions might then become important.

Procedure

A total of 60 rats was used in the experiments. These were divided into 3 main groups, each of which was again divided into 2 sub-groups of 10 animals. As in the preceding experiments, the Lashley jumping apparatus (4) was used, and the animals were required to choose between cards placed in the right and left windows of the screen. One of the cards contained a black circle on a white background, the other a white circle on a black background. One of the cards was locked, resulting in punishment when chosen; the other was free to fall over when the animal jumped against it, admitting the rat to food.

The punishment involved in striking the locked window could be applied consistently to responses to the right (or left) window, regardless of which card was in place; or in connection with either of the cards, regardless of the side on which it appeared; or it could be applied randomly, thus accompanying each card and each position 50 percent of the time.

The most common choice-reactions for the rat are those made on the basis of position or on the basis of the appearance of the cards. A consistent preference for the right or left card, regardless of its appearance, will hereafter be referred to as a *position response;* and a consistent preference for either one of the two cards, regardless of its position, will be referred to as a *discrimination response.*

An acquired response may become modified by changing from one of position to one of discrimination, or from one of discrimination to one of position. We will designate such changes as *shifts* in habits. The habits may also change from right to left position preferences, or vice versa, and from preferences for one of the cards to preferences for the other. Changes of this kind we will hereafter refer to as habit *reversals.*

After the animals had been given preliminary training on the jumping apparatus, the experimental procedure was begun. The training program given to the various groups is shown in Table 1. In all phases of the training 10 trials a day were given, a trial being a jump from the jumping stand to the screen, regardless of whether it was correct or not. Groups 1A and 1B were frustrated at the outset since, no matter what response was given, punishment was received on half of the trials and reward on the other half. This condition soon led animals to refuse to jump. To force a choice compressed air was blown on the animal after it delayed one-half minute. The intensity of the air blast was increased by 3 steps (mild, medium and full) after each half minute of resistance so that after 1½ min. full pressure (10–14 pounds) was used. If the animal still resisted for another 1½ minutes, a pointed stick was used to prod the animal. This procedure was seldom necessary.[2]

[2] Unless some added incentive such as air pressure is used, animals may refuse to jump for hours. Mr. J. B. Klee is using a group of animals in his doctor's problem in which food motivation is used to force a response. Under these conditions a large proportion of animals gradually starve to death because they do not make the 10 required responses in a day's time. Animals which make enough jumps to survive, show characteristics similar to those of rats forced to jump by means of air.

TABLE 1

TRAINING SITUATIONS FOR VARIOUS GROUPS OF RATS

Training Problem	Gr. IA	Gr. IB	Gr. IIA	Gr. IIB	Gr. IIIA	Gr. IIIB
A. *Initial problem* Trained until each rat has made a given response 97 percent of the time in last 160 trials. If rat fails to learn in 200 trials, he is aided by guidance until criterion is reached.	Insoluble problem. Frustrating situation.	Insoluble problem. Frustrating situation.	Position habit. 1st choice made positive side.	Position habit. 1st choice made positive side.	Discr. habit. Black circle on white background.	Discr. habit. Black circle on white background.
B. Habit modified. 200 trials given to change. Rats failing to change in this time are given guidance. Training continued on each rat until 97 percent of last 160 trials are in terms of modified habit.	Must reverse response. Old response punished 100 percent of time.	Must shift to new habit. Old response punished 50 percent of time.	Must reverse position habit. Old response punished 100 percent of time.	Must shift to discr. habit. Old response punished 50 percent of time.	Must reverse discr. habit. Old response punished 100 percent of time.	Must shift to position habit. Old response punished 50 percent of time.
C. Frustration given to all groups. Training continued until 97 percent of last 160 trials are the same.	Insoluble problem.	Insoluble problem.	Insoluble problem.	Insoluble problem.	Insoluble problem.	Insoluble problem.
D. Trained to response differing in kind from one acquired under condition C. 200 trials given to make shift.	Response shifted. 50 percent punishment for old response.	Response shifted. 50 percent punishment for old response.	Response shifted. 50 percent punishment for old response.	Response shifted. 50 percent punishment for old response.	Response shifted. 50 percent punishment for old response.	Response shifted. 50 percent punishment for old response.

As found in the previous investigation (7), the animals placed in the no-solution problem soon developed a consistent manner of responding. Usually it was a consistent position response, but occasionally discrimination responses developed.

To match the responses formed in the no-solution problem, Groups IIA and IIB were rewarded for responding on the basis of position, and Groups IIIA and IIIB were rewarded for choosing the card with the black circle on the white background. The responses formed by these groups were developed under the usual conditions of motivation. To keep the element of frustration at a minimum, the required position habit was made to correspond with the animal's first choice, and the required discrimination was a preference for the black circle on the white background rather than for the white circle on the black background. Maier (5) has previously shown that this discrimination is much more readily formed than the reverse, because the positive card then has a greater total brightness than the negative card.

In Condition B the strengths of the various responses were tested by attempts to modify them. Two kinds of techniques for modifying the behavior patterns were used in each main group. In one technique, the animal was required to make a shift in its response; in the other, it was required to reverse its response. The reversal of a habit involves 100 percent punishment for the old response, and a shift of a habit involves 50 percent punishment. The two techniques thus involved very different distributions of punishment. Because the shift in response was in one direction for Group II and in the opposite direction for Group III, the direction of the shift as a possible contributing factor was controlled.

Since the requirement of a shift in a habit involves 50 percent punishment of the response in progress, as does the no-solution problem, it may be supposed that these conditions are alike. However, the no-solution problem involves a pattern of punishment which is random, whereas the condition for shifting a habit involves 50 percent punishment in accordance with some pattern. In other words, the latter condition involves orderly punishment, providing the animal can experience the order, and the former does not. Whether or not the pattern of punishment is a factor can be determined by comparing the effects of the two conditions of 50 percent punishment.

Condition B thus permits a comparison of the strengths of responses formed under motivating and under frustrating conditions as well as a comparison of different methods of habit alteration.

At the end of Condition B, all rats had again acquired a response which consistently led to food. Those animals which were unable to learn the new habit by themselves were guided by the experimenter. (As previously used (7), this guidance consisted of manually pointing the rat to the positive stimulus, and interfering with its attempts to orient on the basis of its former preference or fixation.)

In Condition C, all animals were subjected to the no-solution problem and received 50 percent punishment in random order for any response they had been making or tended to make. With a response in progress, all animals were now subjected to a period of frustration. This part of the experiment permits a determination of the kind of response to which experienced animals will resort under conditions of stress. Obvious possibilities are: continuation of the response in progress, a return to earlier responses (regression), and the development of entirely new responses.

In Condition D, the presence of fixations formed under Condition C was determined. This was done by attempting to modify the adopted response with a constant pattern of punishment. All animals were required to shift the habit which was in progress at the end of Condition C.

It is apparent that all of the 60 rats could not be tested at the same time. For this reason the whole program was performed 3 times, with each group represented by either 3 or 4 rats. Some animals also completed a stage before the others. This was obviously the case whenever guidance was used to aid some individuals. In order that the experimental conditions be consecutive, each animal was carried through its specified program without intervening periods of waiting until group mates reached the same stage.

An attempt was made to make the groups comparable by so dividing litters among the 6 groups that each group contained the same number of males and females, and the same number of representatives from each litter. In some cases, when litters were small, a deviation from this procedure was necessary. In such cases care was taken to equalize the frustrated *vs.* motivated division (Groups I *versus* II and III), as well as the A and B sub-divisions in the 3 groups. All rats were between 4 and 5 months old when experimentation was begun and 4–5 months older when it ended.

Results

I. *The Responses Formed During The First Training Period (Condition A)*

The results of the initial training situation are such as one might expect on the basis of our previous study. Groups IA and IB, which were confronted with the frustrating no-solution problem, formed position and discrimination responses. The number of rats forming each of these is shown in Table 2. Groups IA and IB were treated alike at this stage and the results are very similar. Taken together, 17 rats of Group I formed position responses and 3 discrimination responses.

Of the 17 rats forming position responses, 9 jumped to the right and 8 to the left. Of the 3 forming discrimination responses, 2 chose

the black circle on the white background and 1 chose the white circle on the black background.

Because all rats were trained until 97 percent of the last 160 trials were of the same nature, the minimum number of trials given to a rat is 160. It will be seen that the average number of trials is somewhat in excess of this number, being 201 for Group IA and 209 for Group B. The ranges in trials for the two groups are also similar despite the fact that individual differences were very marked.

Groups IIA and IIB were rewarded for responding to the right or left position. Since the position response which the animal was required to learn was the one which corresponded to its first choice, we may regard this problem as very simple and almost devoid of frustration. The results given in Table 2 show that most animals

TABLE 2

THE RESPONSES ACQUIRED BY THE VARIOUS GROUPS DURING FIRST TRAINING SITUATIONS (CONDITION A)

Response Formed	Gr. IA	Gr. IB	Gr. IIA	Gr. IIB	Gr. IIIA	Gr. IIIB
Position	9	8	10	10	0	1
Discrimination	1	2	0	0	10	9
Average trials required (including 160 criterion trials)	201.0	209.0	160	161	209.8	190.0*
Range	160–320	160–330	160–160	160–170	180–280	170–210

* This average is based on the 9 rats which learned the discrimination response. The animal which formed the position response did not learn a response in terms of the pattern of reward and punishment and so was perceptually in a different situation.

continued to repeat their first rewarded response throughout the 160 criterion trials. Of the 20 rats which responded in terms of position, 11 chose the right and 9 the left position.

The rats in Groups IIIA and IIIB were rewarded for choosing the card with the white circle on the black background. All but one rat in these two groups formed this response. The number of trials required to reach the criterion was 209.8 for Group IIIA and 190.0 for Group IIIB. The animal which failed to learn the discrimination developed a position response after the first 4 trials, on each of which he received punishment. Through the remainder of the 200 trials which were permitted to learn the discrimination, this animal showed a right position response. To break the position response, the animal was given 30 trials in which it was guided so as to avoid errors. After this aid, it responded to the reward-card during the next 160 trials.

During the initial training period all animals had either developed position or discrimination responses which had been repeated at least 97 percent of the time in 160 trials. In Groups II and III the position and discrimination responses, respectively, had been reinforced by

reward, whereas in Group I the position and discrimination responses had been rewarded and punished an equal number of times. In external appearances the rats of Group I behaved like the rats in either Group II or Group III, although the conditions under which they developed their choices were very different.

The next training condition will reveal whether the similar responses formed under different conditions continue to function in a similar fashion.

II. *The Persistence of Responses Under Two Conditions of Punishment (Condition B)*

In Condition B all rats were given 200 trials in which to alter the previously acquired responses. In each main group half of the rats were placed in a learning situation which required a reversal of their former response, and the other half were placed in a learning situation which required them to shift their response. Those animals which were expected to reverse their response received punishment each time the old response was made (100 percent punishment groups), and those expected to shift their response received punishment on half the trials in which the old response was exhibited (50 percent punishment).[3]

The results of the procedure followed in Condition B are shown in Table 3. Let us first consider the three main groups which formed responses under conditions of either motivation or frustration. It will be seen that rats in Group I (IA and IB), which formed their responses under conditions of frustration, tended to continue their old response throughout the 200 trials given under Condition B. Of the 20 rats, 13 (65 percent) failed to alter their behavior, each showing the old response throughout the 200 trials. Groups II and III, which acquired their responses under conditions of motivation, showed a marked tendency to abandon the former response. Only 10 (25 percent) of the 40 rats failed to make the adjustment. The difference is even more marked when we consider the fact that most of the frustrated rats had developed position responses. Groups IIA and IIB, which also had position responses, showed a marked tendency to abandon this response since only 3 (15 percent) out of 20 failed to make a change.

Comparing Groups II and III, both of which formed responses under conditions of motivation, we find that in Group II 3 of 20 rats

[3] In shifting from a position response to a discrimination response, reward occurs in connection with one of the cards. Since this card is in each position on half the trials, position responses are accompanied by punishment on half the trials. The same is true of shifts from discrimination to position responses. When a response is reversed, punishment obviously accompanies each response made to the old position or the old card, whichever the case might be.

TABLE 3

PERSISTENCE OF RESPONSE THROUGH 200 TRIALS OF FAILURE (CONDITION B)

Treatment of Group in Condition A	Frustrated Groups		Groups Motivated to Learn Position Habit		Groups Motivated to Learn Discrimination Habit	
	IA	IB	IIA	IIB	IIIA	IIIB
Pattern of punishment in Condition B	100%	50%	100%	50%	100%	50%
No. of rats persisting in habit acquired under Condition A. (Rats with fixations)	8	5	3	0	6	1
No. of rats abandoning response for a new one. (Rats without fixations)	2	5	7	10	4*	9
Average no. of trials to abandon first response	18.5 (16–21)	86.2 (30–148)	8.6 (4–21)	34.2 (13–50)	47.8 (40–57)	14.2 (8–20)
Average no. of following trials showing no consistent response, *i.e.*, period between old and new response	16.5 (4–29)	2.6 (1–12)	4.3 (0–9)	10.4 (0–46)	43.3 (0–105)	1.1 (0–3)

* All of these rats failed to learn the problem in terms of motivation which required them to reverse their discriminations, but instead formed position habits.

failed to change and in Group III 7 of 20 rats failed to change. When one considers that the rats of Group II learned a response which corresponded with their natural preference, whereas the rats of Group III learned the more difficult discrimination response (one which was selected by the experimenter), it might be supposed that Group II rats should have a stronger response than Group III rats. If this supposition is valid, it is obvious that the habit's initial strength did not interfere with its relatively easy replacement. If we suppose, however, that in the formation of the more difficult response, some rats were frustrated, then we can explain why Group III showed more persistence than would be expected in terms of learning theory. More evidence on this point will come to light later.

If we now consider the pattern of punishment of the old response, we might expect 50 percent punishment to be less effective for the breaking up of a response than 100 percent punishment. The results, however, demonstrate that the opposite is true. If we combine the *A* sub-divisions of Groups I, II, and III (which received 100 percent punishment), we find that 17 out of 30 rats (56.7 percent) failed to abandon their old responses, whereas in the *B* sub-division of the same groups (which received 50 percent punishment), 6 out of 30 rats (20.0 percent) failed to give up their responses. Other things

being equal, it is obvious that 50 percent punishment induced more rats to abandon their old responses than 100 percent punishment. Since in both A and B divisions some rats were expected to learn position habits and others were expected to learn discrimination habits, the nature of the habit to be learned cannot be regarded as a factor influencing the difference obtained. To explain the ineffectiveness of 100 percent punishment, we must suppose that this pattern of punishment involves frustration, and as we have already seen, frustration produces fixations.

Let us now consider the rats which successfully abandoned the old response. In all cases, excepting Group IIIA, the rats learned the problem which was set for them under Condition B. The *A* sub-groups reversed their responses and the *B* sub-groups shifted from position responses to discriminations or vice versa. The 4 exceptions in Group IIIA all formed position responses rather than reversed discrimination responses. The exceptions will be considered later.

If we examine the number of trials required to break the old response in the animals which failed to fixate, we find that fewer trials are necessary with 100 percent punishment than with 50 percent punishment. These values are given in Table 3, opposite the heading, "Average number of trials to abandon first response." The ranges in scores are shown in parentheses and the numbers of rats on which the values are based are given in the line above. In nearly all cases the *A* and *B* sub-divisions show a marked difference and the ranges in score show little overlapping. Only the 4 rats of Group IIIA, which failed to learn the prescribed problem, are inconsistent with the trend. The observation that 100 percent punishment makes for more rapid destruction of a response than does 50 percent punishment is what may be expected from the principles of reinforcement in learning, but stands in contrast to the number of animals which fixated under the two conditions.

To clarify the issue we must postulate merely that the results follow learning theory to the point where frustration in experience is produced. If the animals do not experience frustration, then 100 percent punishment is highly effective for relearning. If, however, they have a low level of tolerance or do not learn soon enough, then 100 percent punishment becomes frustrating and the resulting fixations appear.

That the abandonment of the old response is readily followed by the learning of a new response which avoids punishment is shown by the values in the last line.

It thus appears that 100 percent punishment causes a ready shift in habit in animals which are relatively immune to frustration; the other animals become frustrated, and punishment has a fixating effect

on the habit. Because the frustrating element is more pronounced in 100 percent punishment than in 50 percent punishment, the former procedure produces the greater number of fixations.

In the case of the 4 exceptional animals in Group IIIA which failed to learn the prescribed problem, we have two sources of frustration. One is the difficulty of learning a discrimination problem, the other is 100 percent punishment. In the 6 group mates either Condition A or 100 percent punishment in Condition B was sufficient to produce frustration and the resulting failure to abandon the old response. That 100 percent punishment is perhaps the main factor is indicated by the fact that the 6 rats which failed to abandon their response had no more difficulty in learning the problem in Condition A than did the 4 rats which did abandon the response. Further, the Group IIIB rats received the same treatment in Condition A and only one of them fixated. The 4 exceptional rats may, therefore, be regarded as being less susceptible to frustration than their 6 group mates, and hence abandoning their old responses as did some rats in each group. However, failure to learn a reversed discrimination over a period of time resulted in frustration of the type encountered by Group I under Condition A. Thus Condition B became frustrating to them and they formed position responses as did most rats of Group I under Condition A. That these animals did not immediately take a position habit after abandoning the old discrimination response is shown by the fact that an average of 43.3 trials was required before they settled down to a position response.

It may be supposed that if more than 200 trials had been given in Condition B, all rats would eventually have learned the problem. An examination of the table, however, reveals that for the 37 rats which abandoned the old response, 148 trials was the maximum score. The 23 rats which did abandon the response continued to make the old response for 200 trials each. In all groups, excepting Group IIIA, 100 percent of the reactions were in terms of the old response. In this group 99.97 percent of the reactions were to the discrimination. It is apparent that rats which abandon the old response and those which do not, form a discontinuous series. This fact was pointed out in our previous studies (**7**, **8**) and it was a basic reason for insisting upon a qualitative distinction between *abnormal fixations* and *habits*.

Fixations are formed primarily under conditions of frustration. A fixated response is characterized by its stereotypy and the difficulty with which it is broken. A habit, however, may be transformed into a fixation if the method of breaking the habit introduces frustration. Individuals differ in the ease with which they are frustrated, so that the same frustrating conditions do not produce fixations in all rats.

The results thus far have shown that fixations are more likely to

take the form of position responses than discrimination responses. This may be due to the fact that position reactions are the more primitive of the two. Either, however, may become fixated.

One rat in Group IB developed a position habit in Condition A, but always jumped so as to use the side of its body rather than its head to strike the card. During Condition B it learned the discrimination, but retained the abortive manner of jumping throughout the period. Mannerisms of this sort are frequently formed and fixated under frustrating conditions. The phenomenon of fixation thus seems to be rather general and is not confined to choice-point behavior.

By the end of 200 trials in Condition B, 37 rats had abandoned the response acquired under Condition A and 23 persisted. Of the 37 which abandoned the old response, 33 acquired the response which was rewarded in Condition B and 4 (of Group IIIA) developed a different response from that demanded by the experimental conditions. To bring all of the rats together, with respect to having a second response which was consistent with the pattern of reward, it was necessary to train further the 23 fixated rats and the 4 which learned the wrong response.[4]

The special training given to the 27 rats involved the guidance procedure in which the animals were manually prevented from practicing the fixated response. As previously found (7), only a few trials of guidance were necessary. After one to 7 trials the rats shifted to the response which was consistent with the pattern of punishment, indicating that the essential association had already been formed but had failed to come to expression because of the dominant fixated response. Only in Group IIIA, in which the discrimination response had to be reversed, was guidance found to be somewhat less effective. The 6 rats, which fixated the first discrimination, formed position habits during the guidance period instead of reversing the discrimination. However, after 22 to 54 guidance trials, all of these rats solved the problem.

By the end of Condition B, all rats had acquired a second response which, when practiced, made it possible for them to avoid punishment and receive food. The second response may, therefore, be regarded as a habit learned under conditions of motivation. This habit was well established, since the criterion was 97 percent correct choices in 160 trials. All rats thus entered Condition C with a response which was different from and dominant over the response acquired during Condition A. If they are now frustrated, will they form fixations? If so, will they fixate the second response, the earlier response, or a new response?

[4] While these 27 rats were given special training, the other 33 passed on to Condition C.

III. *The Effect of Frustrating a Learned Second Response (Condition C)*

The introduction of the no-solution problem resulted in 50 percent punishment for all of the rats, regardless of the nature of their existing response patterns. Since the punishment was not given in any regular order, such as is involved when the problem is changed, we may speak of the pattern of punishment as random rather than orderly as was the case in Condition B. The results of this change in procedure are shown in summary fashion in Table 4. From learning

TABLE 4

EFFECT OF NO-SOLUTION PROBLEM ON SECOND RESPONSE

Group	Character of Second Response	Persisted in Second Response	Showed Variation	Shifted Response
IA	9 Position 1 Discrimination	10	0	0
IB	2 Position 8 Discrimination	9	0	1 (Discrimination to former position)
IIA	10 Position	10	0	0
IIB	10 Discrimination	10	0	0
IIIA	10 Discrimination	6	1 (Throughout period)	3 (1 to new position, 1 to former position, 1 to opposite of former position)
IIIB	10 Position	9	1 (From 30th to 60th trial, then returned to position response)	0
Total	60	54	2	4

theory one would expect such a pattern of punishment to result in a disorganization of the response in progress. The results clearly refute this expectation. Of 60 rats, 54 (90 percent) continued, through 160 trials, to make the response acquired in Condition B. The consistency of the reaction is shown by the fact that 47 of the 54 rats made no deviation and the other 7 made one deviation each in 160 trials.

The tendency to persist in a position response, however, is somewhat greater than the tendency to persist in a discrimination response. Of the 60 rats, 31 had position habits and 29 had discrimination habits when confronted with Condition C. None of the rats with position responses failed to persist, whereas 6 (21 percent) of the rats with discrimination responses failed to persist.

In Table 3 it was found that only 6 of 30 rats (20 percent) per-

sisted in a response under the condition of 50 percent punishment, whereas under Condition C, 90 percent persisted. The difference in the two conditions is the pattern of punishment. Punishment administered on 50 percent of the trials, distributed in a random order, makes for much greater persistence than the same frequency of punishment, distributed in an orderly manner. This shows that rats do not change their behavior because a response fails to work, but rather change from one mode of reaction to another because they have formed new associations. These new associations are obviously formed while the old responses are still being practiced. When punishment is administered in random fashion, new associations cannot be formed.

Only 2 out of 60 rats showed variations in behavior. One rat in Group IIIA gave up its discrimination response after 15 trials, and for the remainder of the test period failed to settle on any specific choice reaction. However, abortive behavior appeared and was consistently used after 95 trials. The other rat was a member of Group IIIB. It abandoned its position habit after 30 trials, showed variable behavior through the next 30 trials, and then returned to the same position reaction. For the next 160 trials its response deviated only once from the position preference.

Four rats shifted their responses while confronted with the no-solution problem. All had discrimination reactions in Condition B and formed position responses during Condition C. They abandoned their discrimination after an average of 53.5 trials (9–117 trials), and after an average of 39 more trials (3–89), reacted in terms of position in more than 95 percent of the next 160 trials.

Since the no-solution problem was a frustrating condition, it might be supposed that the few position responses which did appear duplicated earlier responses and thus support the notion of historical regression as set forth by Hamilton and Krechevsky (**2**), Sanders (**10**), and Mowrer (**9**). An examination of the data shows that one rat from Group IB formed a position habit which corresponded to the one formed in Condition A. The other 3 rats which formed position responses were from Group IIIA. Two of these were among the 4 rats which formed position habits under Condition B and had to be guided before they formed the required reversed discrimination response. In one case the position habit formed under Condition C corresponded with the one existing under Condition B, but in the other case the position adopted was just opposite to the former one. The third rat in Group IIIA had fixated its discrimination in Condition B, received guidance to reverse it, and in Condition C formed a position habit for the first time.

Thus two position responses corresponded to former habits, one

was the reverse, and one was a new habit. These findings give no support to the notion of historical regression, but support the position taken by Maier, Glaser and Klee (7), and Kleemeier (3). Frustration primarily produces fixations, and what response is fixated depends upon a number of conditions. The habit in progress is the most likely possibility. Other things being equal, responses based on position are more probable than others, and it may be argued that they are the more primitive.

Barker, Dembo and Lewin (1) regard regression as the change from complex to more primitive behavior. When the notion of regression is defined in this manner, our data lend support to it.

It now remains to be seen whether the responses expressed during the Condition C were actually fixations or whether they represented the mere persistence of old responses because no alternatives were available in the situation.

IV. *The Measurement of the Fixations Produced by the Previous Period of Frustration (Condition D)*

In Condition C the rats were punished 50 percent of the time in a random fashion. Since any response was punished to the same degree as any other, there was no reason, from the point of view of motivation, why a response should be abandoned. Thus, Condition C determined the responses to which the animals would resort under conditions of frustration. Condition D was designed to determine to what extent the responses adopted by the animals were fixated, and to what extent the responses might be classified as habits. In Condition B it was found that 50 percent punishment in an orderly manner caused rats to abandon their old responses if they were capable. To facilitate changes in behavior with the least amount of added frustration, all animals were, therefore, presented with this favorable condition. To do this, each rat was subjected to a problem which demanded a shift in response (from position to discrimination or discrimination to position, depending on the habit in progress during Condition C). Thus the 50 percent random punishment of Condition C was replaced with 50 percent orderly punishment in Condition D.

The results of this change in the pattern of punishment are shown in Tables 5 and 6. Table 5 records the fixations and Table 6 summarizes the data for animals which shifted their responses in accordance with the pattern of punishment.

The responses in progress at the end of Condition C are shown in the second, third and fourth columns of Table 5. A total of 35 rats had position responses, 24 rats had discrimination responses and one rat showed variable behavior accompanied with marked abortive

TABLE 5

NUMBER OF FIXATIONS PRODUCED BY FRUSTRATION OF SECOND RESPONSES (CONDITION D)

Group	Response Dominant at End of Condition C			No. of Rats Which		Total No. of Fixations
	Position	Discrimination	Variable	Fixate Position	Fixate Discrimination	
IA	9	1	0	6	0	6
IB	3	7	0	2	2	4
IIA	10	0	0	9	0	9
IIB	0	10	0	0	5	5
IIIA	3	6	1	2	1	3
IIIB	10	0	0	7	0	7
Total	35	24	1	26	8	34
Percent Fixations				74.3	33.3	57.6

TABLE 6

ANALYSIS OF LEARNING IN NON-FIXATED RATS

Group	No. Abandoning Divided by No. of Cases		Total No. Solving	Average No. of Trials to		Average No. of Trials to	
	Position for Discrim.	Discrim. for Position		Abandon Position	Adopt New Response	Abandon Discrim.	Adopt New Response
IA	3/9	1/1	4/10	101.0	2.3	119.0	11.0
IB	1/3	5/7	6/10	104.0	16.0	38.0	6.0
IIA	1/10	0/0	1/10	163.0	7.0		
IIB	0/0	5/10	5/10			23.6	20.4
IIIA	1/3	5/6	6/9	121.0	9.0	22.0	14.0
IIIB	3/10	0/0	3/10	91.0	5.7		
Total No. of rats...	9/35	16/24	25/59				
Percent Solving..	25.7	66.7	42.4				
Average				107.1	6.2	33.6	13.3
Range				55–178	0–16	10–119	0–57

jumping. During 200 trials of Condition D, 34 animals (57.6 percent) fixated, or in other words, continued to make the specific response that prevailed at the end of Condition C, despite the fact that a successful response was possible. Of the 35 animals with position habits, 26 (74.3 percent) continued them, and 9 (25.7 percent) learned the required discrimination. Of the 24 rats with discrimination responses, only 8 (33.3 percent) fixated and 16 (66.7 percent) learned the required position habit. The animal which varied its behavior during Condition C continued the same type of variations and the same pattern of abortive jumping. This animal's response may also

be regarded as a fixation, although it has not been included in the above analysis.

It is clear that position responses are more likely to persist than are discrimination responses. This either means that position fixations are the stronger or that the conditions for breaking the position responses were less favorable than those for breaking the discrimination responses. We have already suggested that position responses are favored under conditions of frustration because they are the more primitive. It seems reasonable, therefore, that these responses should be favored in Condition D. Rats with position responses thus tend to continue them and rats with discrimination responses tend to shift to position reactions. The very fact that some discrimination responses, however, fail to yield to the favored positional responses demonstrates the strength and persistence of abnormal fixations.

Since Condition D actually duplicates the situation of Group IB under Condition B, we may compare the number of fixations observed on the two occasions. The 10 rats of Group IB were initially frustrated in Condition A, and then were required to shift their responses in Condition B. Since 5 rats (50 percent) failed to shift, the proportion compares favorably with the 57.6 percent obtained with the group as a whole. These findings indicate that the tendency to fixate a habit in progress is not greatly altered by the existence of a background of previous learning. The failure of past experience to influence fixations again indicates that historical regression does not occur.

In Table 6 we have analyzed separately the learning of animals which shifted from position to discrimination responses, and those which shifted from discrimination to position responses. The second column shows, for each of the groups, the proportion of animals with position habits which changed to discrimination responses. In each case the numerator designates the number changing and the denominator shows the number of rats available for this type of change. The third column gives the same data for changes from discrimination to position responses. The fourth column gives the total proportion of changes in responses.

The numbers of trials required for each type of change are shown in the remaining columns. These learning data are separated into: (1) the number of trials required to abandon the old response, and (2) the additional trials needed before the correct response is adopted and followed for 3 consecutive days.

The averages for the learning of the two kinds of changes are shown in the second last line of the table. These averages are based on the individual scores and are not the grand average of each group average.

Inspection of the table shows that the 9 rats which shifted from position to discrimination responses required an average of 107.1 trials to abandon the position responses, whereas the 16 rats which shifted from discrimination to position responses required only an average of 33.6 trials. This large difference, together with the slight overlapping in scores, is obviously significant. It is clear, therefore, that animals are much more likely to perceive the introduction of the orderly punishment when it is in terms of position than when it is associated with a particular card.

The briefness of the intermediate stage, between abandoning the old response and adopting the new, shows that the new response is practically accepted before the old is given up. In both types of change there were some animals which required no transition period.

In some cases the problem to be learned during Condition D corresponded with one previously learned, and in other cases it was a new problem for the animal. By regrouping the rats it is possible to determine whether this difference in previous experience influenced the tendency for animals to shift their responses in Condition D. We have, therefore, separated the animals on the basis of whether or not they had previously practiced the response which they were expected to acquire in Condition D. For some animals a discrimination response was required and for others, a position response. Table 7 shows the results of this analysis.

TABLE 7

THE INFLUENCE OF PREVIOUSLY PRACTICING A RESPONSE ON TENDENCY TO READOPT THE RESPONSE WHEN IT IS REWARDED

Habit Rewarded in Condition D	Response Previously Exercised			Response Not Previously Exercised		
	No. of Rats Which		Percent Which Fixate	No. of Rats Which		Percent Which Fixate
	Learn	Fixate		Learn	Fixate	
Discrimination Response.....	5	10	66.7	4	16	80.0
Position Response.....	10	7	41.2	6	1	14.3
Total.....	15	17	53.1	10	17	63.0

A total of 32 rats previously had practiced the response which was rewarded in Condition D. Of these, 15 readopted the response and 17 (or 53.1 percent) failed to do so, but instead persisted in the response which prevailed under Condition C. Of the 27 rats which had not previously experienced the response rewarded in Condition D, 10 adjusted to Condition D and 17 (63 percent) failed to adopt the new response. Although the number of failures to shift is slightly greater when the alternative rewarded response has been previously

experienced than when not, the difference is too small to enable one to regard previous experience an important determining factor. The fact that half of the animals persisted in a response adopted during a period of frustration, despite their subsequent exposure to a rewarded and previously acquired alternative, demonstrates the relative unimportance of past experience in the behavior of frustrated animals. The fact that the persisting response led only to punishment further corroborates the view that abnormal fixations are unadaptive in nature and as such must be distinguished from strong habits.

We have already shown that Condition A produced fixations and that these were largely confined to the groups which were placed in a frustrating situation. Condition C placed all groups in the same kind of a frustrating situation and produced fixations in all of the groups. It is of interest now to determine the extent to which previous fixations and the following guidance, which was used to break the fixations, may have influenced the tendency to develop fixations during Condition C.

As already shown in Table 5, Condition D revealed that 34 rats had fixated the response developed in Condition C. A total of 26 rats had previously formed fixations, and 11 (42.3 percent) of these fixated a second time. A total of 33 rats had not previously formed fixations, and of these 23 (69.6 percent) fixated during Condition C. Thus the same condition creates 65 percent more fixations in rats not previously fixated than in rats previously fixated. The group of rats that had not previously fixated included some animals which had been exposed to a frustrating situation in Condition A. These animals had demonstrated a marked tolerance for frustrating conditions, and on this basis might be expected to show a smaller percentage of fixations in Condition C. That the trend is in the reverse direction makes the difference obtained even more significant.

It appears that curing an animal of his first fixation also protects him, to a considerable degree, from forming further fixations. Since we have previously demonstrated (7) that abnormal fixations are forms of compulsive behavior, this finding has important clinical implications.

Discussion of Implications of Major Findings

The experiments reported in this paper present additional data on the nature of *abnormal fixations*. Results supporting the qualitative distinction between fixations and habits and the dependence of fixations on frustration were again obtained. In addition, it was found that learning situations frequently produce frustration and in this manner fixations rather than habits are produced. Frustration

is particularly apparent when attempts are made to modify habits. Changing a habit frequently involves punishment and punishment has a dual function. It serves both as a negative incentive and as a frustrating agent. It is for this reason that the pattern of punishment may influence the breaking of habits.

As a negative incentive, punishment administered 50 percent of the time should be less effective than punishment administered 100 percent of the time. We should, therefore, expect 100 percent punishment of a habit to produce more rapid extinction than 50 percent punishment.

As a frustrating agent 100 percent punishment should produce a frustrating experience more readily than 50 percent punishment. Thus 100 percent punishment should create fixations to the greater degree and thus prevent the habit from being broken.

The results showed that 100 percent punishment either produced a rapid alteration of the habit or failed entirely to alter the habit; whereas 50 percent punishment altered habits more slowly but more surely. This is exactly what one should expect if one recognizes the dual properties of punishment.

That 50 percent punishment given in an orderly fashion is less frustrating than when given in a random fashion demonstrates that habit modification takes place *after* an alternative is found. The animal does not abandon an unsuccessful habit and then find another. This interpretation is further supported by the fact that very few trials were required to reach a high criterion of mastery once the old response was given up. These findings are in distinct opposition to trial and error learning theory which postulates that responses must be rewarded before they are adopted. Actually the association between a correct response and its consequences can be formed before the response is actually tried out.

It was further found that the breaking of fixations is best accomplished by preventing the response from coming to expression. Habits are broken when their expression fails to be followed by reinforcement or when their expression is followed by punishment. In either case the habit must come to expression if it is to be weakened.

Once a fixation is broken, an individual is somewhat less susceptible to the formation of further fixations. Thus specific modifications of fixations are not likely to lead to substitute fixations. The breaking of habits does not reduce the tendency to form further habits.

The above findings have important educational and clinical implications. To produce the best learning conditions, frustrating agents must be avoided. Because punishment is a possible frustrating factor, it cannot be treated as a mere negative incentive. When an undesirable and persistent response must be broken either in

school or in the clinic, it is well first to determine whether the response is a strong habit or a fixation. The methods for breaking habits are entirely ineffective for breaking fixations.

Since fixations arise from frustration and are different from habits formed under conditions of motivation, it follows that frustration and motivation are different instigators of action. Recently Maier (6) has applied this distinction to social movements and presented reasons for recognizing two qualitatively different types of movements.

Summary

The purpose of the present study was to analyze further the nature of abnormal fixations and to investigate in particular the relationship between the pattern of punishment and fixations.

Three groups of animals were used and each group was presented with 4 experimental stages or training conditions. The Lashley jumping apparatus was used in all the training conditions. Animals were required to jump to one of two windows in which cards were placed. When animals resisted making a choice, they were forced to do so by means of a blast of air presented from behind.

Condition A.—In the first training condition the rats of Group I were presented with an insoluble problem. They received punishment (striking a locked card) on half their trials and reward (striking unlatched window and landing on food platform) on the other half, no matter how they chose. In this situation all animals either developed a response based upon position or a discrimination response.

The animals of Group II were trained to develop a position response by being rewarded for this type of choice and punished for any other. Similarly, the rats of Group III were trained to develop a discrimination response by being rewarded for choosing a card with a black circle.

The first training condition thus developed both position and discrimination responses but part of each of these responses was developed under conditions of frustration, and part under conditions of motivation. The frequency with which each response was practiced was the same for all groups (160 repetitions).

Condition B.—The next stage of the experiment was concerned with a modification of the response acquired in the previous training condition. Half of the animals in each group were punished on 50 percent of the trials in which the previously acquired response was expressed, and the other half were punished on 100 percent of the trials in which this response occurred. This was done by requiring animals with position responses either to shift to discrimination responses (50 percent punishment) or to reverse their position prefer-

ence (100 percent punishment), and animals with discrimination responses either to shift to position responses (50 percent punishment) or to reverse their discrimination responses (100 percent punishment). A total of 200 trials was permitted for learning the new response.

As previously reported, the responses developed under conditions of frustration were less subject to modification than those developed under conditions of motivation. In addition, it was found that in all groups the 100 percent punishment resulted in definitely fewer response modifications than did the 50 percent punishment (43.3 percent as compared to 80.0 percent). The 100 percent punishment tended either to produce rapid habit modification or a fixation of the habit so punished.

At the end of the 200 trials the animals which failed to learn were given manual guidance. This procedure was particularly effective for breaking fixated responses.

Condition C.—The third experimental stage was a frustrating condition for all animals in that they received punishment on half of their responses. The punishment was distributed in random order so that there was no escape from the pattern of punishment. This procedure was continued for 160 trials.

Under these conditions 54 of the 60 rats continued to respond in terms of the habit last developed. Only 7 out of 8,640 of their trials were of a different nature. The results clearly show that under these conditions animals persist in a going response. They did not return to an earlier mode of behaving (*i.e.*, show regression), despite the fact that each rat had had a well developed previous response.

Condition D.—In the final experimental stage the rats were required to change their last mode of response. Animals with position responses were exposed to a discrimination problem and those with discrimination responses were exposed to a position response problem. In each case the 50 percent random punishment was changed, therefore, to 50 percent orderly punishment. The orderly punishment offered 50 percent punishment for the response in progress, but complete escape from punishment if the proper shift in behavior was made. This change in the distribution of punishment, being most conducive to learning, constituted a test for the determination of the number of fixations produced by the preceding stage. This procedure was continued for 200 trials.

The results showed that 57.6 percent of the 60 rats had developed fixations. Because the second period of frustration produced about the same proportion of fixations as the first, it was concluded that a background of previous habits does not significantly alter an animal's tendency to fixate the habit in progress.

In some cases the response required of the animals corresponded

with a response previously practiced, and in others the response required was a new one. This difference in previous experience, however, had no significant effect on the results. When an animal has formed a fixation, this fixation dominates the behavior and a previously learned alternative cannot compete.

The results were analyzed to determine whether a previous history of fixations predisposed rats to form further fixations. This analysis revealed that rats without a history of fixations actually showed 65 percent more fixations than animals which had previous histories of fixation.

The discussion treats the implication of the findings to psychological theory. Applications to education and clinical psychology are also suggested.

(Manuscript received November 16, 1942)

References

1. Barker, R., Dembo, T., & Lewin, K. Frustration and regression. An experiment with young children. *Univ. Ia. Stud. Child Welf.*, 1941, **18**, No. 1, pp. 314.
2. Hamilton, J. A., & Krechevsky, I. Studies in the effect of shock upon behavior plasticity in the rat. *J. comp. Psychol.*, 1933, **16**, 237–253.
3. Kleemeier, R. W. Fixation and regression in the rat. *Psychol., Monogr.*, 1942, **54**, 1–34.
4. Lashley, K. S. The mechanism of vision. I. A method for rapid analysis of pattern vision in the rat. *J. genet. Psychol.*, 1930, **37**, 453–460.
5. Maier, N. R. F. Qualitative differences in the learning of rats in a discrimination situation. *J. comp. Psychol.*, 1939, **27**, 289–341.
6. Maier, N. R. F. The role of frustration in social movements. *Psychol. Rev.*, 1942, **49**, 586–599.
7. Maier, N. R. F., Glaser, N. M., & Klee, J. B. Studies of abnormal behavior in the rat. III. The development of behavior fixations through frustration. *J. exp. Psychol.*, 1940, **26**, 521–546.
8. Maier, N. R. F., & Klee, J. B. Studies of abnormal behavior in the rat. VII. The permanent nature of abnormal fixations and their relation to convulsive tendencies. *J. exp. Psychol.*, 1941, **29**, 380–389.
9. Mowrer, O. H. An experimental analogue of "regression" with incidental observations on "reaction formation." *J. abn. soc. Psychol.*, 1940, **35**, 56–87.
10. Sanders, M. J. An experimental demonstration of regression in the rat. *J. exp. Psychol.*, 1937, **21**, 493–510.

AN ANALYSIS OF ERRORS MADE IN THE LEARNING OF PROSE MATERIALS[1]

BY CHARLES N. COFER

The George Washington University

Introduction

Systematic theoretical formulations of the process of learning as it occurs in the learning of lists of nonsense syllables and lists of meaningful words have recently been presented by Hull and his coworkers (**7**) and by E. J. Gibson (**5**). Although these discussions have increased our understanding of learning in these situations, there has been little work done in the analysis of the learning process in the case of connected meaningful material—prose and poetry. Indeed, Hull has said that the introduction of nonsense syllables was a notable technical advance ". . . because they present the subject with material to which, as remote stimuli (*i.e.*, as stimulus traces) he has relatively few reaction potentials already connected, certainly nothing like so many as in the case with isolated words and, even more, connected discourse" (**7**, p. 14). While it is undoubtedly true that the complexities involved in the theoretical analysis of learning of connected discourse are great and that our ignorance concerning many of the questions which it involves is considerable, it seems important to make an attempt at least to suggest some of the important aspects of the problem so that further research may be directed to study them. This is the purpose of the present paper.

One of the most fruitful methods for the study of learning, which has advanced our knowledge of maze learning, serial and paired associate learning of syllable and word lists, etc., is that of the analysis of errors. Thus, it has seemed appropriate to the writer to use this approach in the attempt to analyze the learning process in the case of prose material. This may be done with at least two objects in view: (1) the discovery of factors important in the learning of connected discourse itself; and (2) the discovery of factors already known to operate in simpler verbal learning situations through the direct comparison of the errors on the two types of material.

In serial or paired associates learning of discrete word or nonsense syllable lists, the errors are relatively easy to designate and to classify. Consequently, it is possible to subject them to quantitative treatment

[1] This paper was presented in part at the meetings of the Eastern Psychological Association, April 18, 1941, at Brooklyn, New York (**3**). The data were obtained in 1938–1939 at Brown University, under circumstances described elsewhere (**2**).

and analysis. Unfortunately, this is not the case with errors made in learning connected materials. In the learning of a passage of prose, for example, the errors may involve a single word, a phrase, a part of a sentence, a whole sentence, etc., and the evaluation of these different types of errors is difficult. Hence, the analysis of such data must, at least at present, be qualitative and illustrative, rather than quantitative and inclusive.

Many studies have been made in which prose or poetry has been used as material. In very few instances, however, have records of the trial by trial course of the learning been secured. And in only one instance known to the writer has any effort been made to study the errors in such trials. Oldfield and Zangwill (11) conducted such a study, utilising as materials prose passages which were written so as to be ungrammatical, inconsistent, or normal, *i.e.*, without wilful distortion. Their interest, however, was in the determination of the effects of the abnormal passages on the frequency of various classes of errors. Their analysis was not as specifically directed to the study of ordinary prose as seems necessary to the writer.

Method

The data obtained in a previous study (1, 2) were utilised for the error analysis of the present one. They consisted of complete records of the recitations made on each trial by each of 24 *S*s while learning 16 passages of prose (for the shortest passage length only 14 *S*s were used). The essential aspect of the method was that the *S* read the passage aloud once and then attempted to recite it without reference to the text; this constituted one trial, and all other trials were carried out in the same manner. The *S*s learned 8 passages word for word (verbatim) and 8 for substance, four lengths of passages being available for each method. The criterion of mastery was two successive errorless recitations. The bulk of this report will be concerned with evidence obtained for verbatim learning, although substance learning will also be discussed briefly.

General Classes of Errors

Four classes of errors may be described: (1) word-order errors (2) omissions (3) added material (4) substitutive errors. Each of these will be described below with illustrations, and later an attempt will be made to interpret at least a portion of each type in terms of factors which seem to contribute to its occurrence. Illustrations are given from one or two passages learned verbatim; these illustrations may be considered typical of errors occurring in all the passages so learned.

1. *Word-Order Errors.*—This was a frequent type of error in which a word, phrase, part of a sentence, or, in a few instances, a whole sentence was misplaced in the recitation as compared to the textual word sequence. Illustrations of this type of error follow. In passage 5V,[2] one sentence read: "However, the boy was revived by a drug

[2] Certain of the characteristics of the passages may be found by looking in 2, Table I. The numbers and letters used here to designate each passage refer to the passages so listed in the table. The passages are presented in full in 1.

his father owned." This word-order was modified by 10 of the 24 *Ss* in these ways: (a) making the father an active agent in reviving the boy, *e.g.*, "boy's father revived him with a drug," [3] 4 of the 10 subjects; (b) changes in the position of 'father' with respect to 'owned,' *e.g.*, "drug owned by his father," or "a drug of his father's," 5 of the 10 *Ss*. The next sentence in this passage read: "The wolf was unwary and slept and its tail was pulled out." 16 of the 17 *Ss* who changed this word order did so through the modification of the voice of the last verb, *e.g.*, 'pulled out its tail,' or some minor variant thereof. These types of changes, in which passive verbs were made active, or the form of a possessive was changed, or in which the function of a word in a sentence was changed in some way were very frequent instances of word-order change.

Another type of word-order change which occurred occasionally was that in which a part of one sentence was placed in another sentence instead of in its proper textual locus. One illustration will be given. In passage 5V two consecutive sentences read: "But the boy refused because without his tail the wolf was powerless. The boy became a great hero since, by this refusal, he protected his tribe from the wolf." The latter sentence was modified to read (the italicized part actually belonging to the former sentence): "The boy became a hero *because without his tail the wolf was powerless.* He had protected his tribe by taking the wolf's tail."

2. *Omissions.*—A frequent error was that in which a word, phrase or occasionally a sentence of the text was omitted in the recitation. Some of these occurred in connection with other types of errors such that for the recitation to be grammatically correct certain textual words had to be omitted (*vide supra*). Others, however, appeared when this was not the case, and the discussion will be confined to such instances.

An example of the omission of one word may be found in the data for passage 7V, one sentence of which read: "The Deer's son was killed by some Indians for food." Five of the 24 *Ss* omitted the word 'some' in one or more trials. The omission of a phrase may be illustrated by the recitations of four of the *Ss* who omitted 'the death of' from the following sentence of passage 7V, "The Deer was sad and resolved to avenge the death of his son." The omission of an entire sentence was relatively infrequent; when it occurred, however, it was often found that several of the *Ss* omitted the same sentence on one or more trials. Thus, the middle sentence of the following three (passage 7V) was omitted for one or more trials by 14 of the 24 *Ss*:

[3] The complicated nature of the errors in prose learning is well illustrated here. The words 'was' and 'by' are omitted in this recitation as they have to be if the re-wording is to be good English.

"She screamed and awakened four hunters. They attacked the Deer but he escaped. They followed him and gained on him, so the Deer climbed a tall tree."

3. *Added Material.*—This type of error, in which words, phrases, or, rarely, whole sentences were recited by the *S* although there was no corresponding material in the text, was found chiefly in the initial trials. One illustration will perhaps be sufficient. In passage IV one sentence read: "He awoke and searched for it." One subject added to this sentence ". . . and found it missing." Most of this class of errors were of this type, although there were some instances in which the added material was necessary because of other changes previously introduced.

4. *Substitutive Errors.*—These errors consisted of the substitution of other words for the words of the passage but without essential change in the meaning of the passage. Thus, these errors may be considered to be synonyms of the correct words. This type of error included the simple substitution of a synonym for a word of the text and the substitution of a phrase for a word or words in the text; the phrase might as a unit be considered a 'synonym' for the textual word or words, even though none of the constituent words of the phrase was synonymous with the textual word or words. Examples of the substitutive type of error were very numerous: pronouns were often substituted for a proper noun or other noun and conversely; verb forms were often changed, *e.g.*, a verb in the present tense might be changed to past tense; synonyms or near synonyms were often given, *e.g.*, the use of 'decided' instead of 'resolved' as was done by 12 of the 24 *S*s in passage 7V, the use of 'built' for 'made' in the phrase 'made a fire' as was done by 21 of the 24 *S*s in passage 11V. An instance of the substitution of a phrase for other words is the following (passage 7V): "The hunters chased him" for "They followed him"; chased and followed are not precisely synonymous, but in the context of the passage either would seem to be correct.

Analysis of Errors Without Regard for Class of Error

Certain other points concerning the errors can best be made without consideration for the classes of errors previously described. These points are: (1) similarity in *S*s' errors and place of error; (2) the relative susceptibility to error of 'major' and 'minor' words; (3) persistence of errors; (4) oscillations in recall. These will be described in turn below and their importance discussed in the next section.[4]

[4] It is important, where possible, to analyze errors to determine serial position effects. Unfortunately, in constructing the passages used in this study, the author made no attempt to equalize the difficulty of the sentences in various serial positions. Hence, no comments on serial position effects will be made in this paper. However, see Jersild (8).

1. *Similarity in Ss' Errors and Place of Error.*—A striking aspect of the errors appearing in the recitations in this study is the similarity in the kind of error made by different *Ss* and the tendency for them to make their errors at the same place in a sentence. These tendencies have been amply illustrated already. Thus it was mentioned above that 16 of 17 *Ss* modified the voice of a single verb (similarity in place of error); many of these used exactly the same wording in modifying the voice. As already indicated, 21 *Ss* used the same synonym in changing the wording of a phrase in passage 11V. There were of course many instances of errors made by only one *S*, and in many cases also only one *S* erred at a given place in a sentence. The tendency to similarity in these two respects was very pronounced, however; other illustrations may be found above, and scores of instances could be cited.

2. *The Relative Susceptibility to Error of 'Major' and 'Minor' Words.*—It has previously been suggested (**2**) that 'significant' words in a passage are more easily learned than the passage as a whole. This was seen in the comparison of scores (trials to learn and success per trial) for these words and for the total passage. A more detailed analysis seems to confirm this finding. The results from two sentences will serve to illustrate the point. One sentence of passage 5V read: "Then the wolf became angry, because the boy was playing very roughly." The major or significant words, essential to the meaning of the sentence, were thought to be: *wolf*, *became*, *angry*, *boy*, *was playing*, *roughly;* the minor words were thought to be: *then*, *the*, *because*, *the*, *very*. Fifteen *Ss* made errors on the major words at some point in the course of learning. All 24 *Ss* made errors on the minor words. Most of the errors on the major words consisted of the substitution of *played* for *was playing* (11 *Ss*). The others were the substitution of *grew* or *got* for *became*, *got* for *was playing*, *rough* for *roughly*. Errors on the minor words consisted either in the omission of a word or in the substitution of an equivalent for it. Nineteen *Ss* omitted *then* on some trial and 8 used *but* or *so* in place of it. Thirteen omitted *very*, and 7 used *too* or *so* for *very*. Three cases omitted *because* or substituted for it the words *and* or *for*.

Another sentence, in passage 7V, read: "But one woman was awake." The major words here were thought to be *woman* and *was awake*. The minor words were *but* and *one*. Eight *Ss* made no errors on this sentence, and two omitted it entirely on the first trial but made no error thereafter. For the other 14 cases, there were 15 instances of error on the minor words and 8 on the major words. Most of the errors on minor words were the omission of *but* or the substitution of *a* for *one* (14 cases). Half of the major errors were the substitution of *awakened* for *awake*.

It is clearly suggested by these illustrative data that the minor words lead to a disproportionate amount of error as compared with the major words. It should be noted further that the latter are relatively less often omitted than the minor words; the difference between the two would be more marked if minor variations in word form had been counted as correct. This evidence suggests a possible explanation for certain results reported by Newman (**10**). He found that retroactive inhibition was little present following either a period of sleep or of waking activity in the case of the essential items of the passage, but that it was present to a greater extent after waking activity than after sleep in the case of the non-essential items. Retention was better under all circumstances for the essential than for the non-essential items of the passage. It may be that the non-essential material was less well learned than the important items, which would seem to be a reasonable hypothesis to account for Newman's findings (*cf.* McGeoch, **9**).

3. *Persistence of Errors.*—A striking characteristic of the recitations is that for a given *S* many errors tended to recur throughout the series of trials. Once an error occurred, in other words, it was likely to be repeated. An attempt was made to express this tendency numerically. For each *S* who made an error on more than one trial a tabulation was made of the number of such trials which contained one or more errors which were *identical.* If an *S* made errors on four trials, and on two of these the same error or errors were made, a score of 2 was tabulated. This was then converted into the percentage of the total number of error-trials in which a given error occurred. In the hypothetical case above, the percentage would be 50.

The nature of the data makes even such a gross analysis difficult. For this purpose, an arbitrary procedure was used in determining what would be the error unit, and this was based largely on ease of analysis. Thus, in some cases the errors tabulated were single words, in others phrases, etc. Again for the sake of convenience, only one entry was made for a given *S* per sentence. If two errors persisted in his recitation of a sentence, the one persisting the longer time was used in this analysis; if they persisted an equal number of trials, the number of trials on which one of them occurred was counted. This analysis is by no means precise, but the results are suggestive, and it is probable that the tendency is underestimated by the arbitrary procedures that were adopted.

In Table I, data from two passages are presented. It is seen that for the *S*s who erred on more than one trial on a given sentence the percentage of error trials in which errors persisted is rather high, the median percentage being 62–63.

4. *Oscillations in Recall.*—It has been found (*cf.*, *e.g.*, Hovland, **6**)

TABLE I

DATA SHOWING THE PERSISTENCE OF ERRORS THROUGH SEVERAL TRIALS
(METHOD OF ANALYSIS DESCRIBED IN TEXT)

The data are for two passages learned verbatim. 24 *Ss* learned each passage, but only those having more than one error trial are included in the table.

Passage 5V			Passage 7V		
Sentence	No. of Cases	Ave. Percent of Error Trials in Which Errors Were Repeated	Sentence	No. of Cases	Ave. Percent of Error Trials in Which Errors Were Repeated
1	1	00.0	1	8	40.1
2	21	83.7	2	16	62.5
3	19	77.6	3	23	67.8
4	20	59.4	4	6	69.0
5	17	79.8	5	10	53.4
6	22	64.2	6	19	52.3
7	21	49.1	7	24	51.2
8	23	64.2	8	19	62.8
			9	24	67.4

that a fruitful method for dealing with data on verbal learning is to determine the 'oscillations at the threshold of recall,' *i.e.*, the span in trials between the initial successful or correct response and the last incorrect response. The data for 8 passages learned verbatim have been analyzed in this way. For each sentence of each passage the average number of trials up to the first perfect recitation of that sentence and up to and including the last recitation on which there was an error was tabulated. (Because of difficulties above indicated, no attempt was made to deal with individual words or errors; the sentence was taken as the unit.) Since position effects cannot be discussed meaningfully here (*vide supra*), only the averages for the total passages of the different lengths will be described. Four passage lengths were available—25, 50, 100, and 150 words. Two passages of each length were learned verbatim, so that the averages below are based on data for two passages at each length learned by all the *Ss* (N was 14 for each passage of 25 words and was 24 for all other passages.) The average number of trials up to the initial perfect recitation of the sentences for the passages of 25, 50, 100, and 150 words, respectively, was 0.536, 2.28, 3.03, 3.75. The average number of trials up to and including the last failure for the four lengths was, respectively, 0.608, 2.62, 3.94, 5.21. The average span, derived from these figures, for the four lengths was, respectively, 0.072, 0.34, 0.91, 1.46.[5] These data indicate that, as in the case of nonsense syllable

[5] If an *S* made a perfect recitation on his first trial, the number zero was entered to indicate the number of trials up to the initial success. This is not accurate, of course, but is justifiable in view of the fact that our interest is in comparing the passages of different lengths on this measure. In some cases, particularly with the long passages, some sentences were never learned. This probably decreases the obtained differences among the passages, since oscillations would probably have arisen had these subjects continued to a point of complete mastery.

lists, oscillations in recall occur with prose material, and the data further agree with those presented by Hovland (6) for the serial learning of lists of 8, 11, and 14 syllables in showing increased oscillation as length of material increases.

Discussion

Most of the discussions concerning the learning of connected meaningful material raise the problem of why it is that such material is more easily and rapidly learned than is nonsense material. Various hypotheses have been suggested to account for this difference. Because the evidence presented in this paper does not bear directly on this problem, it will not be taken up here.

Perhaps the most general statement that can be made concerning the data presented above is that the errors made in prose learning appear, for the most part, to arise as a consequence of the *S*'s previous habits, *i.e.*, through transfer of training. In the case of errors, of course, this is negative transfer. The evidence will now be summarized in respect to this point of view, and a few other comments will also be made.

Of the four classes of errors two seem to relate most specifically to negative transfer: word-order errors and substitutive errors; added material will also be discussed in this connection. In the case of the analysis of errors without respect to class of errors, two aspects seem related to this transfer: similarity in *Ss*' errors and place of error and the relative susceptibility to error of major and minor words.

A reasonable hypothesis concerning this negative transfer would seem to be that it arises in two ways: (1) From direct interference with the *Ss*' habitual modes of verbal expression, and (2) from the *Ss*' having learned on numerous previous occasions *techniques of learning* which, applied to a new situation, may be essential to learning but may also account for some of the errors made during learning.

1. *Direct Interference.*—This factor would seem to operate most clearly in the cases of word-order errors, substitutive errors, and similarity in *Ss*' errors and place of errors. It may be involved elsewhere, also, just as other factors may affect the three aspects of errors just enumerated.

In word-order errors, direct interference may be supposed to operate if we assume that the *Ss* with whom we deal have developed habitual modes of verbal expression which, in some instances, would be different from the form employed in the material being learned. Thus, it may be, and it seems reasonable to expect, that the active voice is more commonly employed than the passive in ordinary discourse. If this is true, then the very common tendency of the *Ss* to

introduce word-order changes when the text contains a verb in passive voice may perhaps be attributed to the superior strength of the commonly employed habit. The same type of factor could be supposed to operate in other word-order changes, such as modification of the form of a possessive, or the alteration of the functions of words in a sentence. Other word-order alterations undoubtedly involve additional factors.

Substitutive errors would likewise seem explicable, at least in part, in terms of direct interference. Recent work has shown that generalization of conditioning can occur to synonyms (and other relatives) of a conditioned word (*cf.* 4 for a fuller account of this phenomenon and some of its implications). If we make the assumption that words which have to one another a relation of synonymity are on a generalization gradient, and further, if we assume that the reinforcement of one of them will also increase the 'reaction potentials' (*cf.* 4) of the others, then we have a possibility for explaining some of the substitutive errors. If, for example, before the *S* reads a passage he has extant a gradient on which are words of varying reaction potential, and the word which he *reads* has an initially low reaction potential, the increment thus given to *another* word on the gradient may be sufficient to increase its reaction potential to a point at which it and *not* the correct word would be recited on that trial. Hence, the *S* would err. It should be emphasized also that other factors than synonymity such as homophonymity, quasi-homophonymity and numerous other relationships might operate in the same way (*cf.* 4). What variables are important in causing the different pre-experimental reaction potentials of the several words are not known at present; one of them may be 'familiarity,' but the causative factors are undoubtedly complex and furthermore probably vary from *S* to *S* in the light of particular language biographies.

The relationship of the similarity of *S*s' errors and place of error to direct interference arises logically out of the previous comments if we assume that the *S*s *share* or *have in common* many of these reaction tendencies. One instance suggests that this is true at least in some cases and will illustrate this point and also what has been said above. It was mentioned before that in passage 11V occurred the phrase 'made a fire' and that 21 of the 24 *S*s modified this at some time or other to 'built a fire.' Several of the *S*s commented spontaneously after the passage had been learned that according to their usage they always *built* a fire; they had never made one.

2. *Techniques of Learning.*—This factor is less clearly shown by the data than the one previously described. It seems necessary and fruitful to postulate it, however, at least in connection with the contrast in error susceptibility between major and minor words.

It is widely recognized that in work with the human *S* it is frequently necessary to disguise the nature of the experiment so that the *S* will not 'instruct himself to react' in ways either compatible or incompatible with the purpose of the experiment. In learning experiments of the type herein discussed the usual intention is that learning should go on in as 'natural' a fashion as possible, and hence no attempt is made to control this factor. The answer to the question as to why minor words are more susceptible to error than major words may lie in the description of the way the *S* goes about memorizing the material in question. (Factors such as familiarity cannot be ignored here, but they will operate in conjunction with the tendency being suggested.) We have little evidence on this question in relation to connected meaningful material. It seems possible, however, that the subjects may first learn the important words, the ones that 'carry the meaning of the passage,' before memorizing the less important words and that this would lead to more rapid learning and to overlearning of the major words. (This might account also, in part, for some word-order errors, etc.) The assumption would be that this is a *learned method* which the *S* has learned through experience to use as a helpful or perhaps a necessary procedure in the execution of many learning tasks.[6]

In an attempt to explore this matter a little further, the writer has secured verbal reports from one *S*, who was asked to tell what methods were employed as she attempted to learn prose materials verbatim and for substance. It was the testimony of this *S* that her first attempt was to learn certain words, the important ones, whether she was learning the material verbatim or for substance (for ideas, for essential facts). Hence, while data from one case are hardly conclusive and while the method is not above reproach, this type of activity on the part of the *S* may well account for some of the results previously described.

The case of added material seems even more complex than the types of errors already discussed. It would seem likely to arise from either of the types of transfer outlined above. It might occur as a consequence of associations stimulated by the reading of the text. Again, it may represent a general tendency, which *S*s have learned, to supplement material they specifically recall; this might be especially true when it is evident that they have not recalled in entirety the material they have read. Probably both factors operate conjointly.

[6] Some writers in describing similar phenomena, *e.g.*, grouping, imply that this is a 'natural' procedure, *i.e.*, apparently it would be resorted to irrespective of whether the *S* had ever employed such methods in his previous history. I see no reason for making this assumption (*cf.*, in this connection, McGeoch, 9). The problem of how learning methods are developed in the child is one deserving of more attention than it has received.

It should be emphasized that the above speculations are tentative and require much further investigation before they can be considered established. Because there is little information with which to begin, there has been no attempt at rigor in the above formulations.

Omissions and Oscillations in Recall.—These are discussed together because they both seem related to findings of nonsense syllable experiments. Some, at least, of the omissions may be interpreted as place-skipping errors, due to the formation of remote associations. (Some word-order changes might also be so construed.) This is a frequent interpretation of place-skipping errors in nonsense syllable learning, but there is other and independent evidence of the formation of remote associations in this case. The writer is aware of no other conclusive evidence as to the formation of remote associations in the learning of connected materials, so that further comment seems unjustified. The discovery of oscillations in recall in the learning of prose may suggest that since they also occur in nonsense syllable learning (*cf.* **6**) the same mechanisms may operate in both cases.

Perseverative Errors.—The large number of these errors is difficult to interpret, as it may be a function of the particular method employed (*cf.* **2**) and not a general phenomenon. These errors seem to occur with greater frequency in this study than in the usual nonsense syllable investigation.

Substance Learning

There seems little reason to postulate any factors not already discussed in the description of substance learning, as it was carried out in this study (**2**). There was some evidence in the full verbal reports of the *S* mentioned above that under the instruction to learn for substance she concentrated even more on the major words than when she was learning material verbatim. So far as the shape of the learning curves (**2**) and the types of errors are concerned, however, there seems to be little difference in the two procedures. The major variation between them, as here studied, seems to lie solely in the criteria of mastery; in substance learning many responses are credited and others (*e.g.*, word-order errors) are not counted wrong which would be called incorrect if the verbatim criterion were being applied. The low correlations ordinarily obtained between these two types of performances would not seem explicable on the basis of the evidence of this study. It may be, however, that some *S*s are better able to use one method, others another, which might account for the low relationships usually reported. This would seem to be a fruitful avenue for investigation.

It may be briefly noted in connection with substance learning that

there were, in the initial trials, evidences of the same types of changes in the recitations that are ordinarily reported only for long-time retention studies. Such changes as 'condensation,' 'rationalization,' etc., were observed to occur over the very short retention interval between the reading of a passage and its recitation.

Summary

Various aspects of errors made in the learning of prose materials are described and discussed. An attempt is made to interpret them in terms of negative transfer arising from the *Ss*' previous language habits and the methods they employ in learning.

(Manuscript received November 24, 1942)

References

1. Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. Ph.D. thesis, Brown University, 1940.
2. Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. *Amer. J. Psychol.*, 1941, **54**, 1–20.
3. Cofer, C. N. An analysis of errors made in the learning of prose material. *Psychol. Bull.*, 1941, **38**, 726.
4. Cofer, C. N., & Foley, J. P., Jr. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychol. Rev.*, 1942, **49**, 513–540.
5. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
6. Hovland, C. I. Experimental studies in rote learning theory. VII. Distribution of practice with varying lengths of list. *J. exp. Psychol.*, 1940, **27**, 271–284.
7. Hull, C. L., *et al.* *Mathematico-deductive theory of rote learning—a study in scientific methodology.* New Haven: Yale Univ. Press, 1940. Pp. 329.
8. Jersild, A. Primacy, recency, frequency, and vividness. *J. exp. Psychol.*, 1929, **12**, 58–70.
9. McGeoch, J. A. *The psychology of human learning: An introduction.* New York: Longmans, Green & Co., 1942. Pp. xvii + 633.
10. Newman, E. B. Forgetting of meaningful material during sleep and waking. *Amer. J. Psychol.*, 1939, **52**, 65–71.
11. Oldfield, R. C., & Zangwill, O. L. The acquisition of verbal repetition habits. *Brit. J. Psychol.*, 1938, **29**, 12–26.

STUDIES IN DECISION

II. An Empirical Test of a Quantitative Theory of Decision

BY LEON FESTINGER

State University of Iowa *

The present article is primarily concerned with the fitting of theoretically predicted functions to empirical data of decision-time and relative frequency of judgment. The theory on the basis of which the predictions have been made is presented in its basic form by Cartwright (1). The situation of making a decision between two things is analyzed into the various forces acting on the subject toward choosing one or the other alternative. In the case of a psychophysical judgment, the magnitudes of the forces toward each of the alternatives depend upon the subjective probability that a response is the correct one.

Cartwright and Festinger (3) make several important extensions of the theory. The authors add the concept of a restraining force against making a decision, this force requiring the difference in forces toward the alternatives to reach a certain magnitude before a decision will be reached. Further steps in the advancement of the theory are made by stating precise assumptions as to the way in which these driving and restraining forces fluctuate in time. These distributions are all assumed to be normal. By deriving the distributions resulting from pairings of these forces, the authors were able to state exact equations between the independent variables of decision-time and relative frequency of judgment, and the dependent variables of magnitude of mean difference between driving forces and magnitude of mean restraining force.

These equations show relative frequency of judgment in an ogival relationship with magnitude of mean difference between driving forces, with the 50 percent point of relative frequency at the point of zero difference in driving force. The precision of this ogive increases as the magnitude of mean restraining force increases.

These equations show that decision-time reaches a peak at the point of zero difference in driving forces and drops off symmetrically on both sides along the abscissa of magnitude of mean difference in

* From the Child Welfare Research Station.

driving force. The peak of the decision-time curve is a positive function of the magnitude of mean restraining force.

The theory offers the possibility of quantitative measurement of magnitude of force. This, together with the fact that independent measures of the variables concerned are obtainable,[1] makes an empirical test of these precise, quantitative predictions feasible.

Procedure

The exact experimental procedure used is described in detail in another article (4). The method of constant stimuli was used in a two-category psychophysical experiment. The subjects were asked to judge whether one line was longer or shorter than another line. Forty judgments were made at each of 15 positions of the variable stimulus. Three measures were recorded: the response made, the time that elapsed from the presentation of the stimuli to the giving of the response, and the subject's rating of confidence in his judgment. Five subjects were used in each of four conditions:

(a) Usual instructions.

(b) Accuracy instructions. These instructions were assumed to increase the restraining force against making a decision.

(c) Speed instructions. These were assumed to decrease the restraining force against making a decision.

(d) 'Constant error' instructions. These instructions were assumed to make the restraining force against giving one response (*e.g.*, longer) greater than against giving the other response (*e.g.*, shorter).

Results

Figs. 1 and 2 show the theoretical curves of decision-time and of relative frequency fitted to the corresponding empirical data for two subjects under all four conditions. Figs. 3, 4, and 5 give a representative example of the data from each of the other subjects. On all of these figures the left hand ordinate represents relative frequency in percent (*R.F.*); the right hand ordinate represents decision-time (*D.T.*) in seconds; the top abscissa represents confidence (*C*, negative representing shorter and positive representing longer); and the lower abscissa represents the z scale (difference in driving force) as coordinate with the Confidence scale.

The smoothed curves are in all cases the theoretical curves. The ogive is the theoretical curve of relative frequency and the other is the theoretical curve of decision-time. The open circles represent the empirical decision-time values and the closed circles represent the empirical relative frequency values. In each figure there are 15 pairs of empirical values (one of each pair being decision-time and the other

[1] Decision-time and relative frequency of judgment can obviously be measured empirically. Mean differences in driving forces, since they depend upon differences in subjective probability of either response being correct, are measured by the subject's ratings of confidence as to the correctness of his judgment. The measurement of the magnitude of the mean restraining force is accomplished in the fitting of the theoretical functions to the data, although increases and decreases in restraining force are also defined in terms of the experimental instructions to the subject.

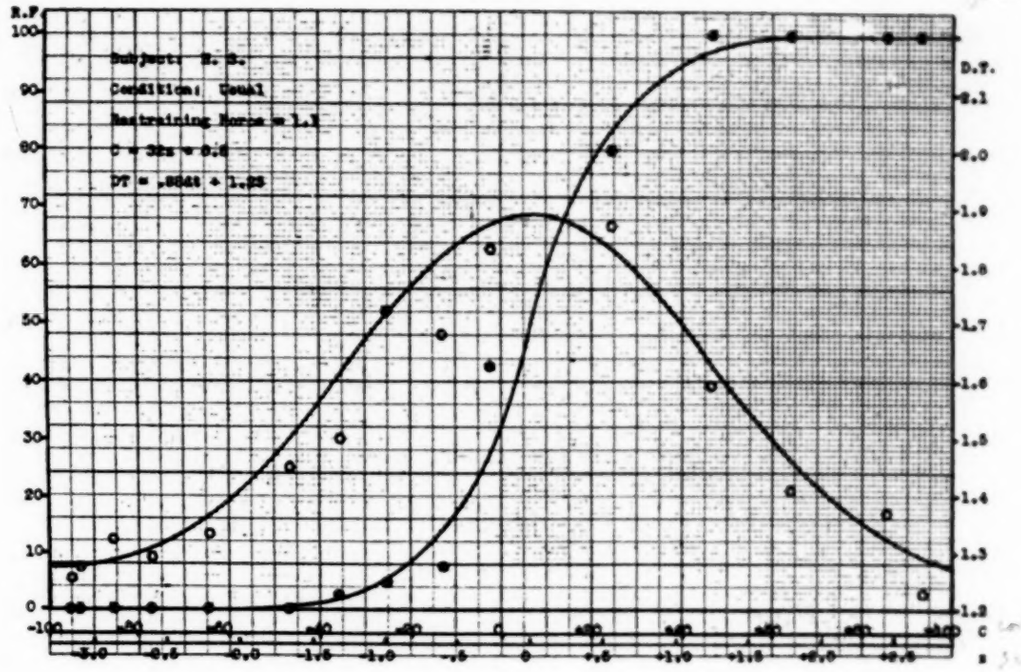

FIG. 1a. Theoretical decision-time and relative frequency curves fitted to the data for Subject H. S. under 'usual' instructions

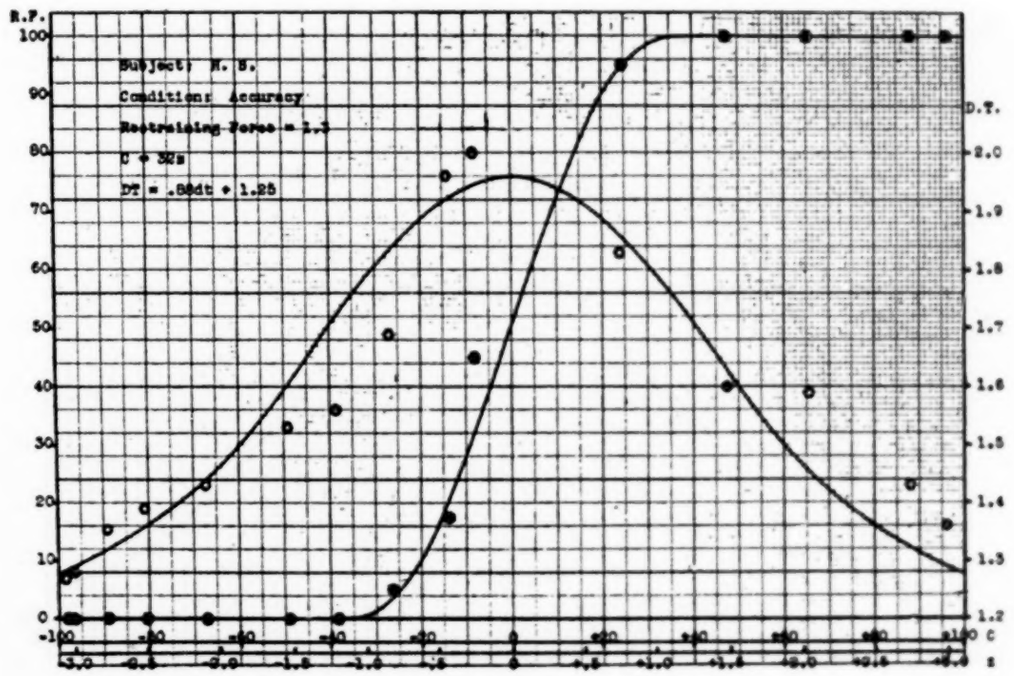

FIG. 1b. Theoretical decision-time and relative frequency curves fitted to the data for Subject H. S. under 'accuracy' instructions

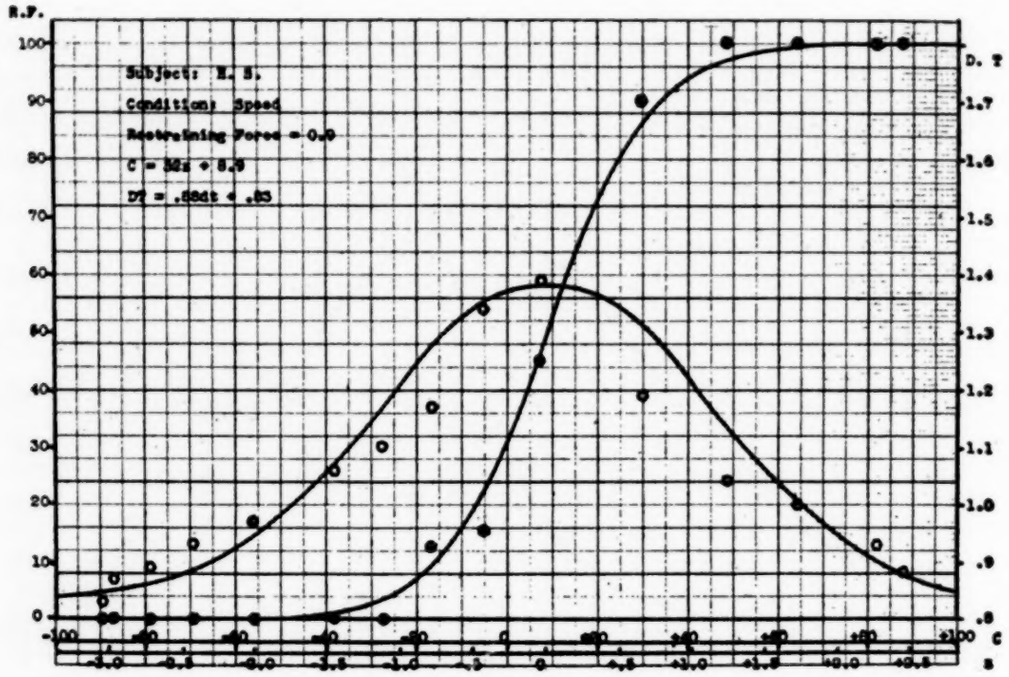

FIG. 1c. Theoretical decision-time and relative frequency curves fitted to the data for Subject H. S. under 'speed' instructions

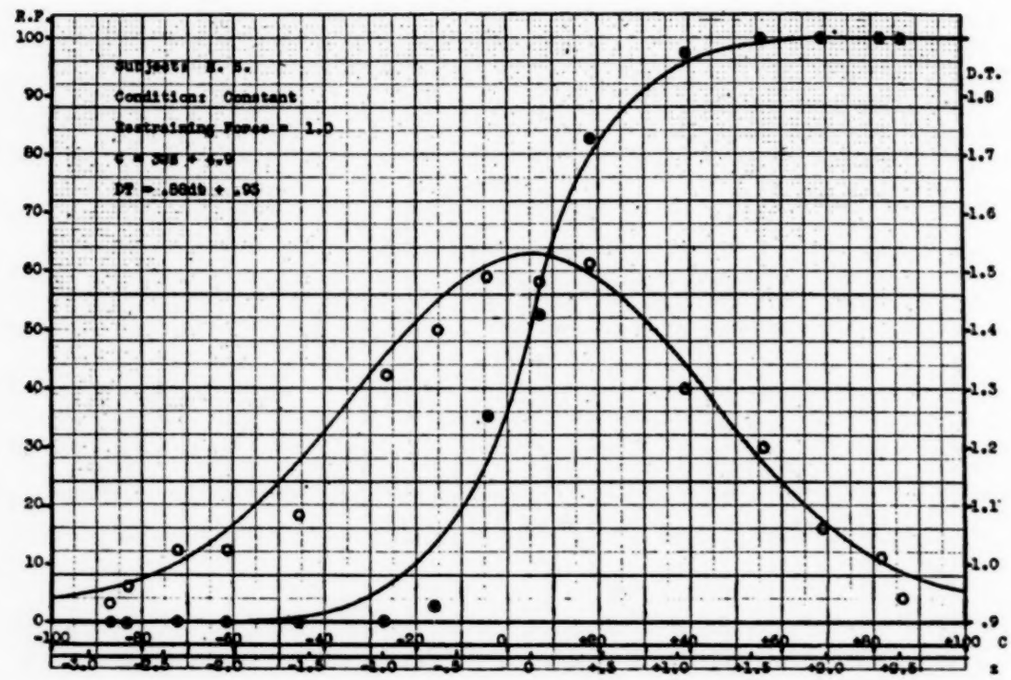

FIG. 1d. Theoretical decision-time and relative frequency curves fitted to the data for Subject H. S. under 'constant error' instructions

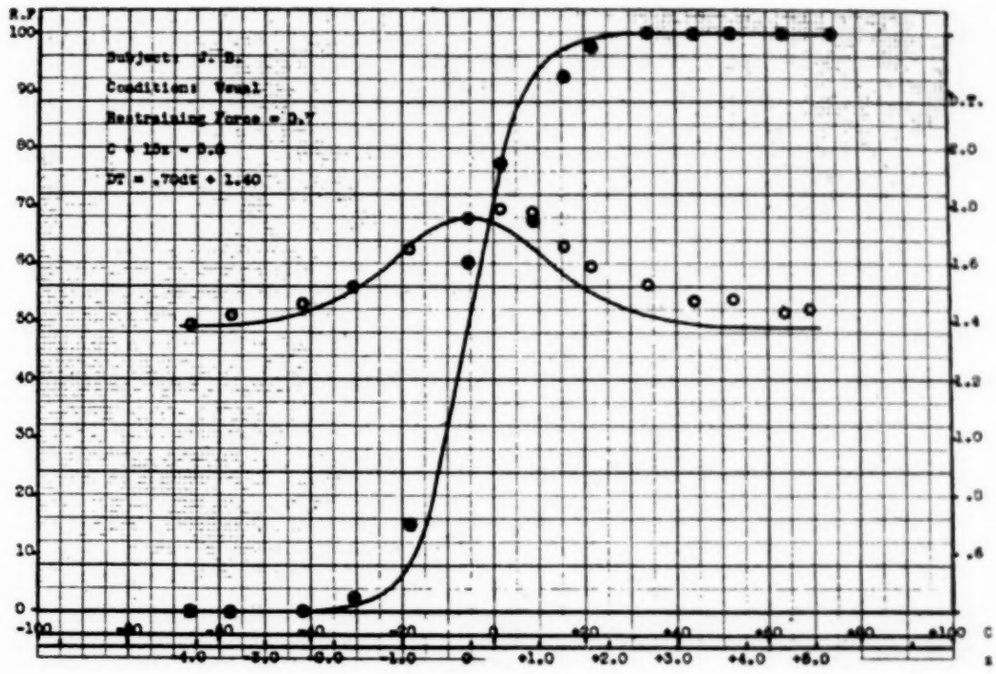

FIG. 2a. Theoretical decision-time and relative frequency curves fitted to the data for Subject J. B. under 'usual' instructions

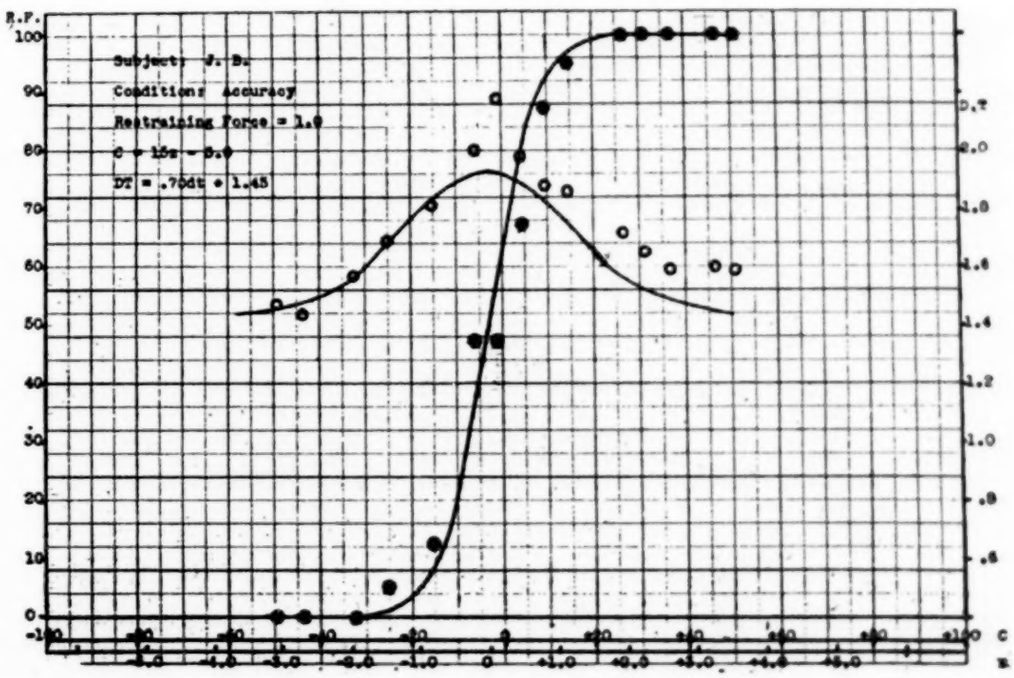

FIG. 2b. Theoretical decision-time and relative frequency curves fitted to the data for Subject J. B. under 'accuracy' instructions

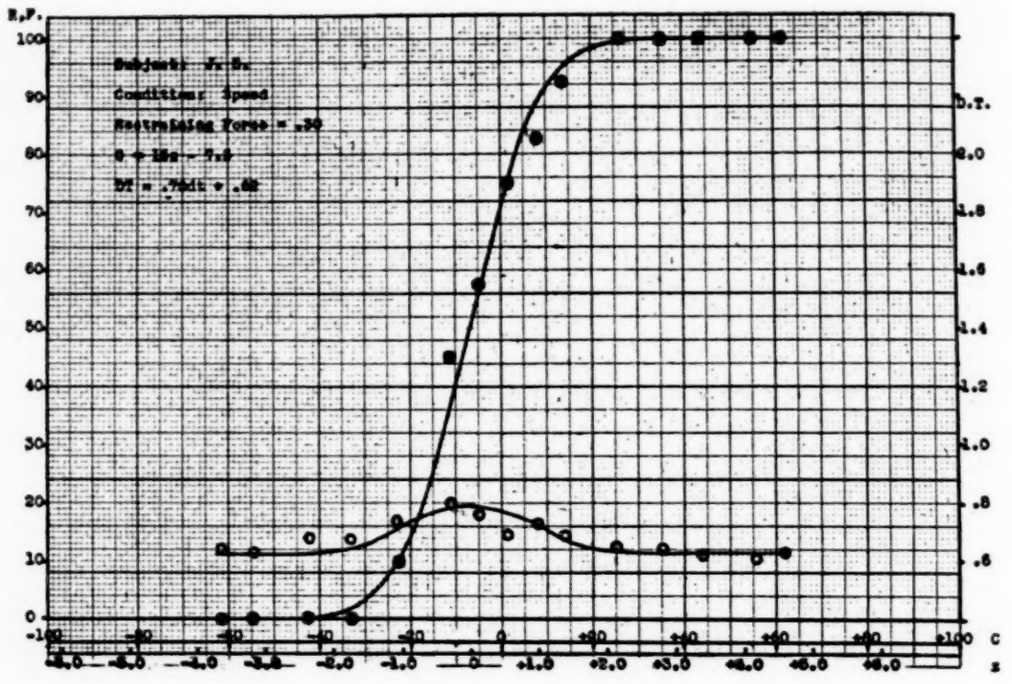

FIG. 2c. Theoretical decision-time and relative frequency curves fitted to the data for Subject J. B. under 'speed' instructions

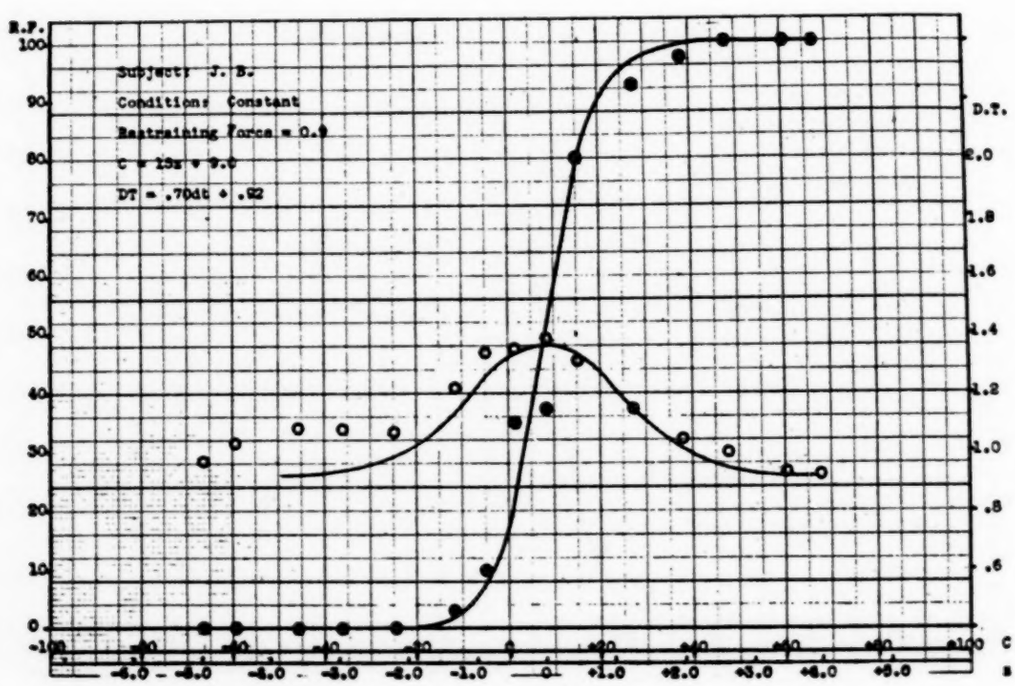

FIG. 2d. Theoretical decision-time and relative frequency curves fitted to the data for Subject J. B. under 'constant error' instructions

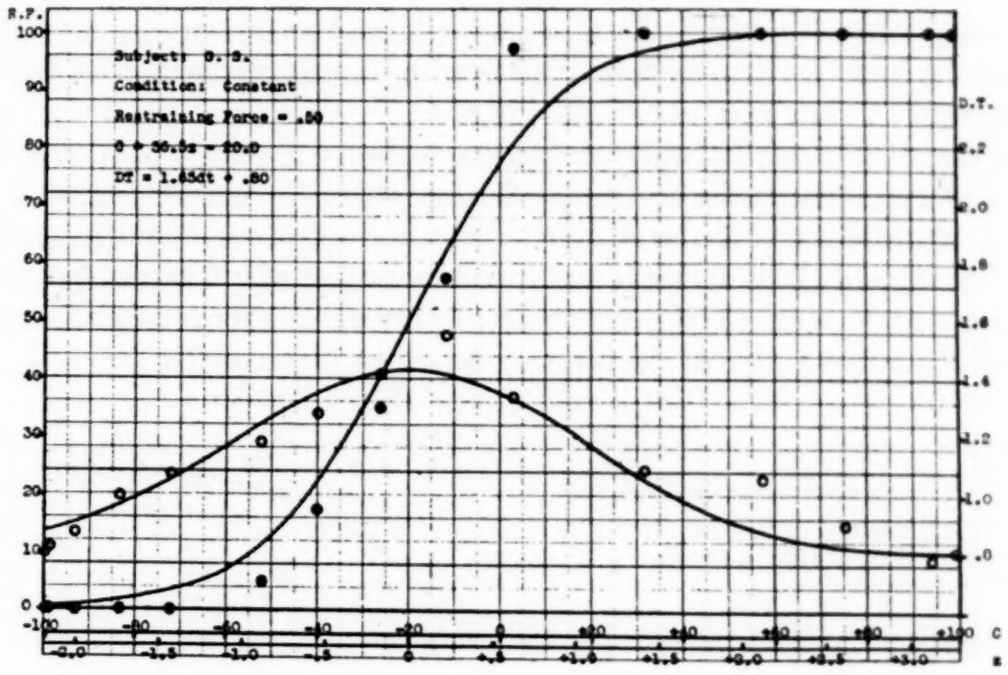

FIG. 3. Theoretical decision-time and relative frequency curves fitted to the data for Subject G. S. under 'constant error' instructions

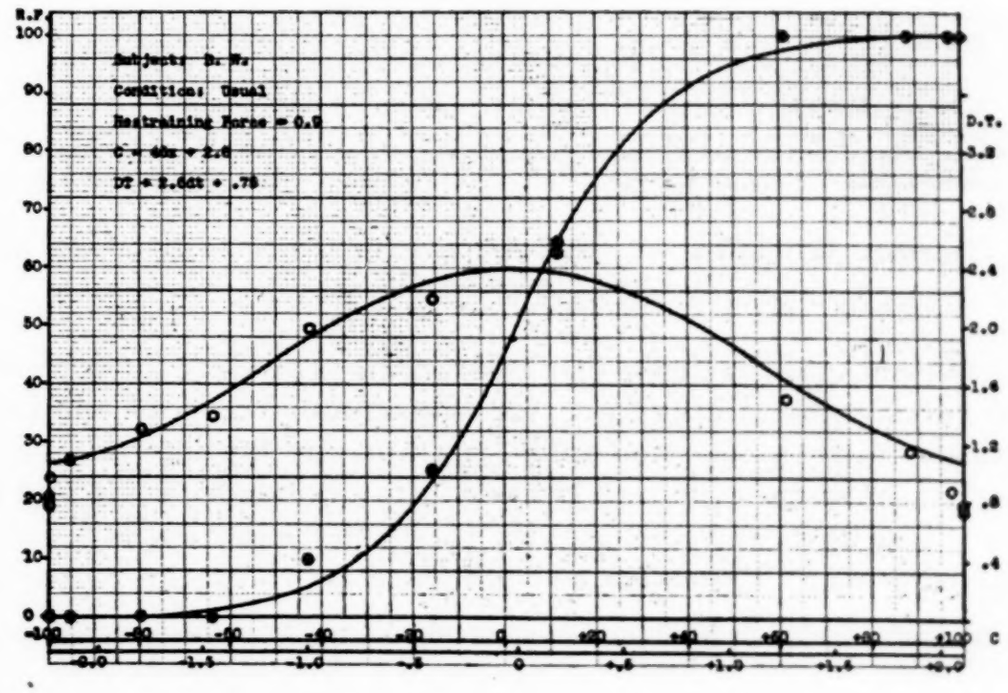

FIG. 4. Theoretical decision-time and relative frequency curves fitted to the data for Subject B. W. under 'usual' instructions

relative frequency) corresponding to the 15 positions of the variable stimulus. These empirical values are plotted at that point along the abscissa which corresponds to the average confidence rating for the judgments at that stimulus position.[2]

In the upper left hand corner of each figure are indicated the subject, the experimental condition, the magnitude of restraining force for that condition, calculated by the method explained by Cartwright and Festinger (3), the linear equation between the Confidence scale and the z scale, and the linear equation between theoretical decision-time units and empirical decision-time units in seconds.

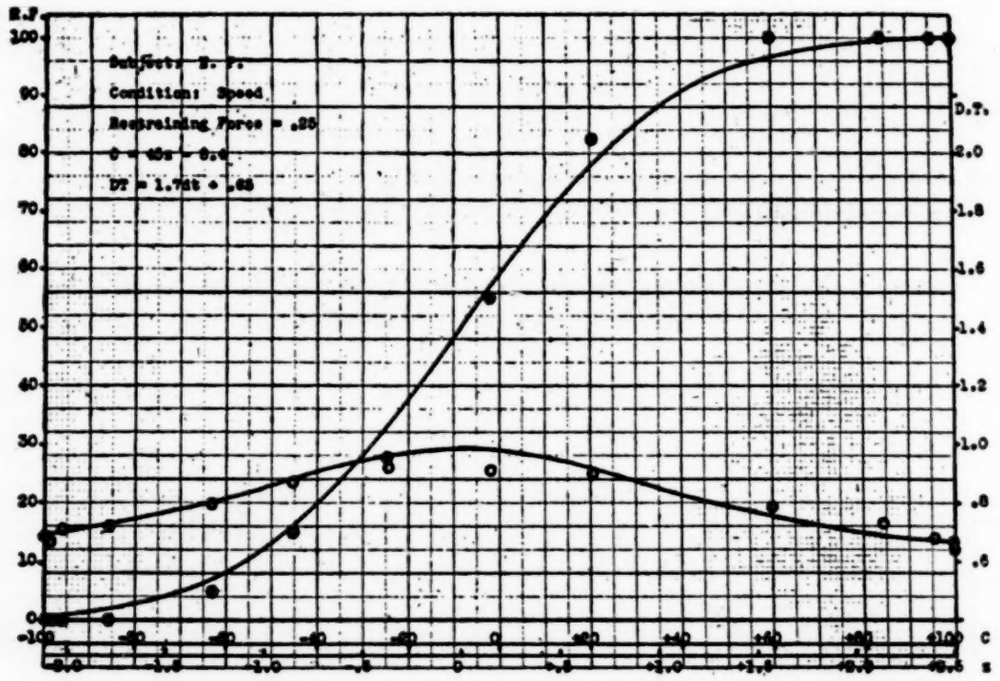

FIG. 5. Theoretical decision-time and relative frequency curves fitted to the data for Subject N. P. under 'speed' instructions

Cartwright and Festinger (3) describe the procedure for determining the value of restraining force, for determining the slope and intercept constants of the decision-time equation, and for determining the slope constant of the equation between the Confidence scale and the z scale. The intercept constant for this latter equation is found in the following way: The value of confidence at which the 50 percent

[2] In another article (4) it has been shown that the relationship between confidence and stimulus difference may best be described as an ogive. Therefore, the smoothed ogive, rather than the empirical values, has been used along the abscissa for plotting the data. Some fluctuation of the standard deviation of the confidence function was present from condition to condition. These changes in the standard deviation were shown to be a random affair. For the purpose of plotting the data in Figs. 1 through 5, the confidence functions for each subject have been corrected to the standard deviation of the confidence function under the 'usual' condition for that subject.

point of relative frequency falls is determined. This value of confidence must be the intercept constant of the equation in question, since the 50 percent point of relative frequency must coincide with z equal to zero.

Figs. 1a, 1b, 1c, and 1d show the data for Subject H. S. under the 'usual,' 'accuracy,' 'speed,' and 'constant' conditions respectively. For all conditions the constant of conversion of C to z is kept at 32.0 and the constant of conversion of DT (decision-time in seconds) to dt (decision-time in theoretical units) is kept at .88.

Figs. 2a, 2b, 2c, and 2d show the same data for Subject J. B. In this case the constants are again identical for all four conditions; namely, 15.0 and .70 respectively.

These values for Subject G. S. (Fig. 3) are 36.5 and 1.65; for Subject B. W. (Fig. 4) they are 46.0 and 2.60; for Subject N. P. (Fig. 5), 43.0 and 1.70. These constants are also identical for all four conditions for these last three subjects.

The problem of empirically 'proving' or 'disproving' a quantitative theory reduces, in psychology as in other sciences, to the question whether or not the equation of a theoretical relationship fits the distribution of empirical data.

There are two general procedures which may be used in fitting theoretical curves to empirical data when conversion or transformation constants are involved. (Such constants are ones which relate empirical units to theoretical units.)

(a) Placing no restrictions (except the general shape of the curve) on the process of fitting and allowing the conversion constants to emerge from this process. The theoretical curves must then be evaluated on the basis of how well they fit the empirical data and how nearly alike are the obtained constants.

(b) Determination of the conversion constants beforehand and then placing the restriction upon the process of curve fitting that these constants must remain unchanged. The theoretical curves must then be evaluated in terms of how well they fit the empirical data.

We have used method (b) in fitting the theoretical functions of decision-time and relative frequency to the empirical data. The values of the conversion constants were determined from the 'usual' condition and then kept unchanged for all other conditions. These constants might also have been estimated from all four conditions together, but it was felt to be a more severe test of the theory to use but one condition for this purpose.

Thus for the 'accuracy,' 'speed,' and 'constant error' conditions, only one factor influencing the theoretical curves was variable; namely, the magnitude of the restraining force.

Let us now examine the graphs of the individual subjects.

For Subject H. S. (Fig. 1) the restraining force values for the 'usual,' 'accuracy,' 'speed,' and 'constant error' conditions are found to be, in order, 1.1, 1.3, 0.9, and 1.0. The increases and decreases in restraining force are all in line with the purposes of the experimental conditions; namely, the value for 'accuracy' is higher than for 'usual;' for 'speed' it is lower than for 'usual;' and for 'constant error' it is higher than for 'speed.' In other words, the experimental conditions produced the desired effect. A glance at the figures will show that the decision-time starts to rise long before the relative frequency curve rises. This is more marked for the higher restraining forces, but clear in all cases. (This result may be observed for all the subjects.) The theoretical curves seem to fit the empirical data rather well for all four conditions.

For Subject J. B. (Fig. 2) the restraining forces for the four conditions are, in order, 0.7, 1.0, 0.3, and 0.9. The restraining force in all four conditions follows the expectation. The curves of both decision-time and relative frequency fit well for all the conditions here. The only important deviations are asymmetry of the empirical decision-time data in the 'usual,' 'accuracy,' and 'constant' conditions. In these cases the empirical data run parallel to, but slightly above, the theoretical curves on one side. This is clearly due to the fact that the base constant is different for 'longer' and 'shorter' with this subject. This difference in base constants is not relevant, however, to the theory.

For Subject G. S. (Fig. 3) the values of restraining force, in order, are 1.2, 1.3, 0.7, and 0.5. For this subject the changes in restraining force are in line with our expectations with the exception of the 'constant error' condition, where the restraining force is lower than in the 'speed' condition. (This is the only instance where the restraining force magnitudes do not follow our expectations.) There does not seem to be any explanation of this deviation except for the possibility that on being told that she had been making a serious error, this subject simply decided to 'give up.' (If the subject does not care whether she is right or wrong, her restraining forces against making a decision are weak.) The fit of the theoretical curves to the data seems fairly good in all but the 'accuracy' condition, where the deviations of the empirical data from the theoretical curve (below on the left side and above on the right side) seem to be due to a slight non-correspondence between the 50 percent relative frequency curve and the peak of decision-time in the empirical data. For this subject the relative frequency curve fits least well for the 'constant' condition.

For Subject B. W. (Fig. 4) the restraining force values, in the same order as above, are 0.9, 1.4, 0.8, and 1.4. While the curves for the 'usual' and 'speed' conditions fit the data very well, for the other two

conditions the decision-time data show a consistent deviation from the theoretical curve. The empirical data at the extremes lie consistently below the theoretical curve. A plausible explanation presents itself if we refer back to the confidence function which for this subject was very steep, much steeper than for any of the other subjects (see Festinger (4)). That is, he tended to have either very little confidence or else complete confidence in his judgments. This means that the Confidence scale was less of a discriminatory instrument for this subject. Since there was a limit of 100 on the scale, the average confidence values at the extremes were pushed toward the center, thus producing the obtained discrepancy between the theoretical curve and the empirical data. As one would expect from this interpretation, this effect is more marked for the higher restraining forces ('accuracy' and 'constant error' conditions) since in such instances the decision-time curve has a steeper slope at the extremes. This effect is still noticeable, however, for the lower restraining forces ('usual' and 'speed' conditions) at the very extremes where there are several decreasing decision-time points, all with an average confidence of 100. The relative frequency curves fit well except for the 'constant' condition.

For Subject N. P. (Fig. 5) the restraining force magnitudes, in order, are 0.5, 1.0, .25, and 0.4, all in line with expectation. The theoretical curves seem to fit the data without unusual deviations for all the conditions.

Further Analysis of 'Constant Error' Condition

In the above discussion we have stated only one number as the value of the restraining force for the 'constant error' condition. Actually there should be two numbers, since the restraining forces are of different magnitudes in the two directions. The values of restraining force, given above for the 'constant error' condition, are the average values. We should, however, be able to determine the restraining force in each direction because of the prediction that the 50 percent point of the relative frequency ogive moves in the direction of the stronger restraining force by a distance along the abscissa equal to one-half the difference between the restraining forces (3). If we determine the average location of the 50 percent point of relative frequency on the Confidence scale for the first three conditions and regard this as the 'normal,' then the difference between this 'norm' and the corresponding location in the 'constant' condition, converted into z units, should be equal to one-half the difference between the restraining forces. The position of the 50 percent point on the Confidence scale is given directly by the intercept constant of the equation between C and z.

For Subject H. S. the average intercept for the first three conditions is + 5.23. There is a shift, then, of − .33 confidence units or − .01 z units in the 'constant' condition. The restraining force against "saying 'shorter'" is 1.01, and in the opposite direction it is .99, a difference which is not very large.

For Subject B. W. the average confidence position of the 50 percent relative frequency point for the first three conditions is − 1.2. This position in the 'constant' condition shifts − 14.6 confidence units or − .32 z units. The restraining force opposite to "saying 'shorter'" is 1.72 and the restraining force in the other direction is 1.08.

For Subject G. S. the position of the 50 percent point is − 14.4 for the first three conditions. The extent of the shift in the 'constant' condition is − 5.6 confidence units or − .15 z units. The restraining force opposite to "saying 'shorter'" is .65, and in the other direction it is .35.

For Subject J. B. the 50 percent point lies at an average position of − 5.1 on the Confidence scale. This point, for the 'constant error' condition, shifts + 14.1 confidence units or + .94 z units. This, if correct, would mean that the restraining force against "saying 'longer'" is 1.8 and the other is zero. (The restraining force can obviously not be below zero.)

For Subject N. P. the average position of the 50 percent point for the first three conditions is + .7. In the 'constant error' condition it shifts + 21.3 confidence units or + .50 z units. In this case again, the estimated difference means that the restraining force against "saying 'longer'" is .80, and the other is zero.

These estimates of the difference between the restraining forces in the 'constant' condition are subject to much error. The intercept constants of the equation between C and z are quite variable from condition to condition. Thus the average of three does not provide a very stable or reliable measure to use as a norm.

All the shifts, however, are in the direction expected from the instructions and serve to support further the theory.

Summary

This is the last of three articles concerned with the quantification of a theory of decision and the empirical test of precise derivations from this theory.

In the first paper (3) a theory of decision which combines a topological and vectorial analysis of a decision situation was elaborated mathematically. By means of this mathematical elaboration a successful quantification of the theory was accomplished. Precise

derivations with regard to decision-time and relative frequency of judgment were stated in a form amenable to empirical testing.

An experiment specifically oriented toward empirically testing some of these derivations was described in the second paper (4). The theoretical predictions were evaluated in the light of their agreement with the experimental results. The data were found to be sufficiently in line with the gross theoretical expectations.

The present paper has presented the detailed data from this experiment to which the theoretically predicted curves of decision-time and relative frequency have been fitted. Thus we have here presented a very rigid, quantitative test of the theory. In view of the fact that the theoretical curves fit the data rather well, the theory should be regarded as corroborated.

The type of theoretical analysis used, in addition to the specific theory of decision, has methodological significance. Irrespective of the school of thought or of the theoretical constructs involved, this type of approach can be used to measure the dynamic factors in a psychological situation. From measures of frequency and decision-time we have been able to infer the magnitudes of the forces operating.

This type of analysis should be applicable to a wide variety of psychological problems.

(Manuscript received December 18, 1942)

Bibliography

1. Cartwright, D. Decision-time in relation to differentiation of the phenomenal field. *Psychol. Rev.*, 1941, **48**, 425–442.
2. Cartwright, D. The relation of decision-time to the categories of response. *Amer. J. Psychol.*, 1941, **54**, 174–196.
3. Cartwright, D., & Festinger, L. A quantitative theory of decision. (To be published in *Psychol. Rev.*)
4. Festinger, L. Studies in Decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *J. exp. Psychol.*, 1943, **32**, 291–306.

PROBLEM SOLUTION BY MONKEYS FOLLOWING BILATERAL REMOVAL OF THE PREFRONTAL AREAS

II. Delayed Reaction Problems Involving Use of the Matching-From-Sample Method.*

BY T. SPAET AND H. F. HARLOW

University of Wisconsin

I. Purpose

The purpose of the following investigation was to study the performance of rhesus monkeys, following bilateral destruction of the prefrontal areas, on problems involving spatial and non-spatial delayed reactions.

II. Subjects

Two rhesus monkeys, numbers 54 and 55, were used as subjects. These animals had been previously tested on a series of discrimination and discrimination-reversal problems following a prolonged taming period. Pertinent data concerning the operations, pretest experience, and results on the discrimination and discrimination-reversal problems were presented in the preceding paper (5).

III. Apparatus

The apparatus used in these tests differed from that used in other matching-from-sample tests made at the Wisconsin Laboratory (3, 12), in the following respects: (1) the top, back and sides of the cage were made of presswood instead of bars or wire, (2) the tray holding the stimulus object did not run on metal tracks but remained fixed throughout the test period, (3) the frame enclosing the one-way vision screen moved in a horizontal rather than in a vertical plane, and (4) the interior of the cage was brightly illuminated at all times by a 60 watt lamp.

Two Union Leader tobacco cans and two small copper oil cans served as stimulus-objects. The essential characteristics of the apparatus and the stimuli are shown in Fig. 1. In the picture the one-way vision screen is drawn back for photographic purposes.

IV. Procedure

Preliminary Procedures.

Prior to the beginning of the delayed reaction tests, the subjects were trained to do the matching-from-sample problem, following the general procedure described by Weinstein (12). In this test the subject was trained to displace first the sample-object in the left part of the test-tray and then to displace the identical one of a pair of choice-objects lying in the right part of the test tray. Responses both to the sample-object and to the correct choice-object were rewarded with food. In keeping with the fundamental matching problem, the nature of the

* This research was supported in part by a grant from the Special Research Fund of the University of Wisconsin for 1941–42.

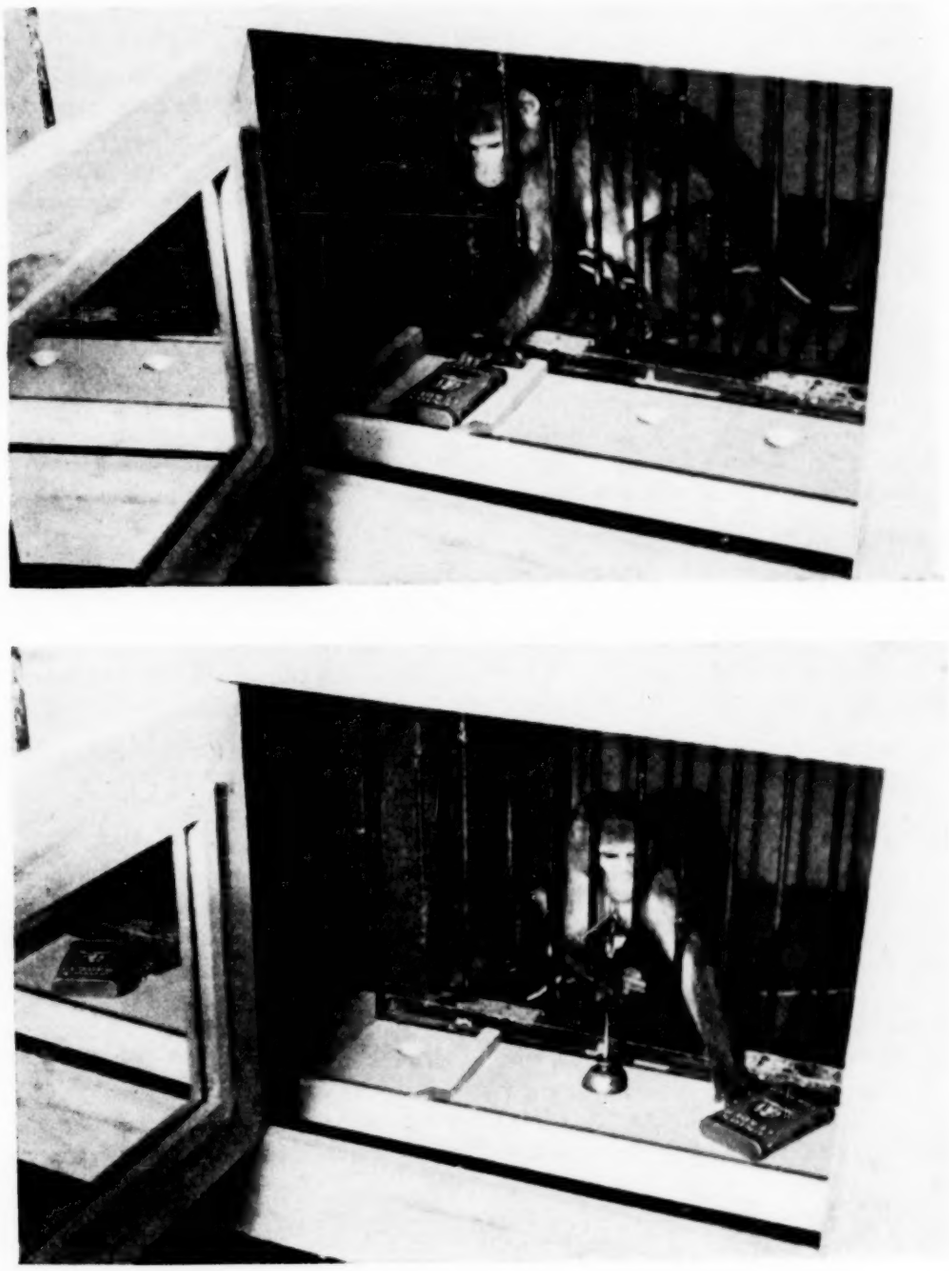

FIG. 1. Solution by Subject 54 of a non-spatial delayed reaction, using the matching-from-sample technique.

Ia. Response to the sample-object.

Ib. Response to the correct choice object following an interposed delay interval.

sample-object, and the nature and position of the correct choice-object was reversed from trial-to-trial in such a manner that the order of presentation could not serve as an adequate cue for problem solution.

General Test Procedures.

The matching-from-sample technique was used throughout the series of delayed reaction tests reported in this paper. The innovations which were introduced were designed to test various aspects of the delayed reaction problem, but in no way altered the basic matching relationship, which demands response to the principle of sameness for the solution of the problem.

A 'correction' method, similar to that described by Krechevsky (9) was used throughout these tests. This technique involved repeated presentation of the same stimulus situation until

a correct response was made. A trial was not considered to be complete until the food reward was obtained; but it was scored as an error if the first response was incorrect.

All trials in all tests in which the period of delay exceeded five seconds were run with a one-way vision screen between the subject and the experimenter, so that the experimenter was never visible to the monkey at the time of choice.

The stimulus-objects were always placed on the test tray in a definite order; the leftmost object was set down first, the middle object next, and the rightmost object last.

Five unmodified matching-from-sample trials were run prior to each day's testing, unless otherwise indicated.

Specific Test Procedures.

Test I. Interrupted matching-from-sample tests.

The procedure of Test I differed from that of the unmodified matching-from-sample test only in the fact that the forward, opaque screen was lowered and left down for a predetermined period of time after the monkey had displaced the sample-object and before he had displaced either choice object. Such a procedure does not involve delayed reaction since the sample-object was still present when differential response to the choice-objects was made.

Test I was divided into the following two steps:

Step 1: 20 trials were run each day. During the first 10 trials the screen was lowered and immediately raised. During the second 10 t.ials the screen was left down for 5 sec. before being raised.

Step 2: 20 trials were run each day. Ten trials in which the screen was left down for 5 sec. were followed by 10 trials in which the screen was left down for 10 sec.

The criterion for solution of this test was nine out of 10 correct responses at each of the interruption intervals of the two steps.

Test II. Spatial delayed reaction.

Test II was so arranged that it could be solved as a spatial delayed reaction. The essential difference in procedure between Test II and Test I was that the sample-object was not on the board when the monkey made the differential response to the choice-objects. The details of procedure are as follows: All three stimulus-objects were appropriately arranged and baited with the forward screen down. The forward screen was raised, the animal was allowed to displace the sample-object, the forward screen was again lowered, *and the sample-object was removed.* After the prescribed period of delay the forward screen was again raised and the monkey allowed to make a differential response to the choice-objects.

Test II was divided into four steps, each step involving two delay intervals which became progressively greater. Ten consecutive trials were run each day at each delay interval and the criterion for success in any step was not less than 90 percent correct response for both delay intervals.

Step 1: 10 zero-sec. delays followed by 10 five-sec. delays.
Step 2: 10 five-sec. delays, followed by 10 ten-sec. delays.
Step 3: 10 ten-sec. delays, followed by 10 twenty-sec. delays.
Step 4: 10 twenty-sec. delays, followed by 10 thirty-sec. delays.

Test III. Non-spatial delayed reaction.

Test III involved non-spatial delayed reactions and could be solved only if the animals remembered the *nature* of the stimulus-object without regard to position. The sample-object and the two choice-objects were never presented simultaneously, with the result that the position of the correct choice-object could not be ascertained until after the period of delay. The following procedure was used: *The sample-object was placed in the appropriate position with no other objects on the test tray.* The forward screen was then raised, and the monkey was allowed to displace the sample-object. The forward screen was now lowered, the choice-objects were arranged in their appropriate positions, and food was put under the correct one. The sample-object was removed, the forward screen was raised, and the animal was allowed to make a choice. The general procedure is illustrated in Fig. 1.

Test III was divided into the following two steps:

Step 1: A five-sec. delay was interposed between presentation of the sample-object and the choice-objects. Twenty trials a day were run until a criterion of 17 errorless trials in a single day's run was attained.

Step 2: 10 five-sec. delay trials and 10 ten-sec. delay trials were run each day. The criterion for solution was 75 percent correct responses in 50 consecutive trials at each of the delay intervals.

Test IV. Non-spatial delayed reaction with an unrewarded sample on half the trials.

The procedure in Test IV was identical with that of Test III save for the fact that in half the trials no food was placed under the sample-object. The trials in which the sample was rewarded and the trials in which the sample was unrewarded were presented in a systematically varied order so arranged that the nature of any trial could not be predicted on the basis of serial cues.

The test was divided into two training steps:

Step 1: Each day 10 five-sec. delay trials were followed by 10 ten-sec. delay trials until a criterion of not less than 75 percent correct responses in 50 consecutive trials of both rewarded and non-rewarded trials had been attained at the shorter delay interval.

Step 2: Subsequently, 10 ten-sec. trials alone were presented each day until the animal attained a score 3 standard deviations beyond chance for any block of 25 or 50 trials. In running these 10-sec. delay trials, the last two of each day's preliminary unmodified matching-from-sample trials were replaced by 2 five-sec. non-spatial delay trials.

Some deviations from the outlined procedure were made in the case of Subject 55. These changes have been indicated in the results.

V. Results

Preliminary Procedures.

The preliminary matching-from-sample learning data of the two subjects are presented in Table O.

TABLE O

Number of Training Trials and Number of Correct Responses in the Last 25 Trials of the Matching-from-Sample Problem

Subject Number	Training Trials	Number Correct Responses in Last Day's Run
54	675	25
55	1100	23

These data show that the number of trials taken to learn by the operated animals is similar to that of normal monkeys (3), (12). It should be noted, however, that there were differences in the training techniques and these may be of significance.

Specific Test Procedures.[1]

Test I. Interrupted matching-from-sample tests.

The learning data for Test I which involved merely adapting the animals to the lowering and raising of a screen between the time of displacement of the sample-object and the selection of a choice-object are given in Table I.

[1] After solution of the matching problem and before the initiation of the specific test procedures reported in this paper, both animals had been tested for delayed reactions, Subject 54 being run for 1125 trials and Subject 55 for 950 trials. Subject 54 learned to solve 15-sec. delays and Subject 55 learned to solve 5-sec. delays to a level of accuracy significantly beyond chance. A one-way vision screen was not used in this preliminary training period.

TABLE I

NUMBER OF TRIALS REQUIRED TO SOLVE THE INTERRUPTED MATCHING-FROM-SAMPLE PROBLEM

Subject Number	Step 1 (0-sec. and 5-sec. Interruptions)	Step 2 (5-sec. and 10-sec. Interruptions)
54	40	20
55	20	20

These data show that the period of interruption occasioned by lowering and raising the forward screen during the matching trials did not produce more than a transient interference in the matching responses of either of the subjects.

Test II. Spatial delayed reactions.

The data of Test II in which spatial delayed reactions were obtained using the matching-from-sample method are presented in Table II.

TABLE II

NUMBER OF TRIALS TAKEN TO SOLVE THE SPATIAL DELAYED REACTION PROBLEM

Subject Number	Step 1	Step 2	Step 3	Step 4
	0-sec. and 5-sec.	5-sec. and 10-sec.	10-sec. and 20-sec.	20-sec. and 30-sec.
54	20	40	160	60
55	480	540	220*	

* Subject 55 failed to attain the criterion at the 20-sec. delay interval of this step.

The results, presented in Table II, show that both of the monkeys with bilateral destruction of the prefrontal areas were capable of solving delayed reaction problems. The performance of one of these animals, number 54, is not inferior to that of normal rhesus monkeys tested under similar circumstances (**12**).

Test III. Non-spatial delayed reaction.

Table III contains the data for the two subjects on Test III. This test could be solved only as a non-spatial delayed reaction since the sample-object and the choice-objects were never presented simultaneously.

These data suggest that rhesus monkeys following recovery from bilateral prefrontal lobectomy, are capable of solving non-spatial as well as spatial delayed reactions. The difficulty of the non-spatial delayed reactions is indicated by the large number of trials taken to learn, by the relatively short intervals of delay successfully bridged, and by the poor performance by Subject 55. The rate of learning by Subject 54 on this uncomplicated test of non-spatial delayed reaction

TABLE III

NUMBER OF TRAINING TRIALS AND PERCENTAGE OF CORRECT RESPONSES IN THE LAST TRIAL GROUP CONSTITUTING THE CRITERIAL INTERVAL FOR LEARNING THE STEPS OF TEST III

Subject Number	Step 1. Training Trials	(5-sec. delay) Percentage Correct Last 20 Trials	Step 2. Training Trials	(5-sec. and 10-sec. Delay) Percentage Correct Last 50 Trials of	
				5-sec. Delay	10-sec. Delay
54	600	85	240	80	76
55	620	85	140*	78	64

* At this point Subject 55 attained the best record he was to make on Test III. At the conclusion of 140 more trials his score was 64 percent correct for the last 50 trials of 5-sec. delays and 60 percent correct for the last 50 trials of 10-sec. delays.

is comparable to that of two unoperated rhesus monkeys similarly tested (**12**).

Test IV. Non-spatial delayed reaction with an unrewarded sample on half the trials.

The data for Test IV are given in Table IV. These data are measures of the performance of the subjects on non-spatial delayed reactions when reinforcement to the primary variable, the sample-object, was withheld on half the trials.

TABLE IV

NUMBER OF TRAINING TRIALS AND PERCENTAGE OF CORRECT RESPONSES IN THE LAST TRIAL GROUP CONSTITUTING THE CRITERIAL INTERVAL FOR LEARNING THE STEPS OF TEST IV

Subject Number	Step 1. Training Trials	(5-sec. and 10-sec. Delays) Percentage Correct Last 50 Trials of		Step 2. Training Trials	(10-sec. Delays) Percentage Correct Last	
		5-sec. Delay	10-sec. Delay		25 Trials	50 Trials
54	520	80 (84)	58 (74)	140	80 (72)	72 (76)*
55	200	76 (88)	60 (68)	120	56 (44)	66 (62)†

Numbers not in parentheses refer to non-rewarded sample trials. Numbers in parentheses refer to rewarded sample trials.

* Throughout the last 100 trials of step 2, Subject 54 consistently maintained a general level of 70 to 80 percent correct responses for both the rewarded and non-rewarded sample trials. This subject was subsequently tested on 100 trials of 20-sec. delayed trials, employing the general technique of Test IV. He made 54 percent correct responses on the 50 non-rewarded sample trials and 62 percent correct responses on the rewarded sample trials.

† Although Subject 55 failed to solve step 2, he did make 24 correct out of the first 25 responses on the rewarded sample and 19 correct out of the first 25 responses on the unrewarded sample and thus failed to attain the criterion by reason of a single error. Subsequently his score fell rather than improved. Eighty five-sec. delay trials were run in the hope that effective testing on his part would be reinstated. Subject 55 responded successfully to 20 of the last 25 non-rewarded sample trials and to 19 of the last 25 rewarded sample trials. Unfortunately, when the delay interval was again increased approximately chance scores returned. The experiment was concluded at this point.

The data of Table IV show that the monkeys attained a level of success on the non-reinforced non-spatial delayed reactions which is similar to the reinforced delayed reaction trials presented in Table III. These data show that the actual food reinforcement does not play an indispensible role in the formation and brief retention of a non-spatial representative factor.

VI. Discussion

Jacobsen and co-workers (**6, 7, 8**) have clearly shown that monkeys following destruction of the frontal association areas suffer a greater deficit in their ability to solve delayed reaction problems than in their ability to form discriminations between stimulus-objects and between stimuli exhibiting marked discrepancies in their figure-ground relationships.

Before evaluating such results it is essential to consider the differences between discrimination problems and delayed reaction problems. Nissen, Riesen and Nowlis (**11**) have suggested two essential differences: delayed reaction learning is characterized by "(1) learning in one *versus* many repetitions of a constant stimulus-reward relationship and (2) learning with signified or 'expected' *versus* actually administered rewards. . . ."

At least one other highly significant difference exists between discrimination learning and delayed reaction learning. In the delayed reaction situation there are *frequent reversals of the correct and incorrect stimuli, i.e.*, the stimulus correct on one trial is frequently incorrect on a succeeding trial. Thus delayed reaction is not a series of discriminations learned in a single trial to a signified reward but is a series of discrimination *reversals* learned in a single trial to a signified reward.

To account for the severity of the loss by monkeys on delayed reaction problems, Jacobsen (**7**, p. 565) has written as follows: ". . . integration over a period of time cannot be carried out effectively. Temporal patterning fails because the subject can no longer remember a *single* (italics ours) experience for even a few seconds in the face of new incoming sensory data." It is obvious that Jacobsen is emphasizing *frequency of repetition* as the factor accounting for the differential susceptibility of delayed reactions and discrimination learning to loss following prefrontal lesions. Furthermore, Jacobsen believes that the primary factor responsible for the failure of one trial learning is the interference of subsequent sensory data.

It is very likely that bilateral prefrontal monkeys are unusually susceptible to disturbances from incoming sensory data. Jacobsen has cited observational data confirming this point and the experiment

by Malmo (**10**) shows that disturbing stimuli may reduce the efficiency of delayed reactions at least during the first few hundred learning or re-learning trials. Malmo (**10**) has gone farther and suggested that retroactive inhibition may account for the rapid extinction of the representative factor, but the design of his experiment is not completely adequate to test the hypothesis advanced and the conclusion is not entirely supported by his data.

There is now a considerable body of data to indicate that, contrary to Jacobsen's suggestion, neither 'temporal patterning' nor ability to 'remember a single experience' are abolished by bilateral destruction of the prefrontal areas. It is very likely that there is loss but that it is only a partial loss.

Finan (**1**) has demonstrated that the prefrontal monkey can learn temporal discriminations. Both Finan (**2**) and Malmo (**10**) have reported delayed reactions learned by monkeys following extensive frontal lesions, and the data of the present paper suggest that both spatial and non-spatial delayed reactions can still be accomplished. Harlow and Dagnon (**5**) demonstrated that bilateral frontal lobe destruction did not prevent either discrimination learning or discrimination reversal from frequently occurring in a *single trial.*

Nissen, Riesen and Nowlis (**11**, p. 382) have ". . . suggested that the loss of delayed response ability after bilateral ablation of the frontal areas should be interpreted . . . as a failure to form associations which involve symbolic responses rather than as a diminution of memory or the retention of associations. The operated animal cannot forget what it has not learned."

This hypothesis, that loss may be occasioned by failure to form associations, receives support from the fact that in all experiments in which monkeys with extensive lesions of the frontal lobes have solved delayed reactions, special procedures have been used to guarantee active response to the constant stimulus-reward relationship. Malmo (**10**) used an unusually intense cue stimulus, a bright light in a dark field. Finan (**2**) preceded each test trial by a reinforced overt response to the cue stimulus. In the present experiment response to the sample-object and reward of the sample-object intensified the relationship for the subsequent delayed reaction.

However, even if we assume that failure to form association is the reason or one of the reasons for the difficulty of the prefrontal monkey in responding adequately in delayed response situations, the problem of the various factors contributing to this failure to form associations still remains.

Finan (**2**) has suggested that ". . . *kind* of reënforcement, rather than *frequency* of reënforcement, offers a possible basis for a differentiating criterion . . ." between the discrimination situation and

the delayed response, and that ". . . one important function of the frontal lobes is the reënforcement of response-tendencies in the absence of any overt or 'natural' reënforcing agency."

There appears to be little question but that the frontal monkeys can, on the other hand, solve delayed reactions in the absence of any overt reinforcing agency. In Malmo's experiment the monkeys apparently did respond adequately to signified rewards. One of Finan's subjects responded adequately to signified rewards after delayed responses had first been established using the technique of a preliminary single reinforcing trial. As the non-spatial delayed responses in the present experiment were run, the response which is retained throughout the delay interval, the selection of the correct choice-object, must be formed on the basis of a signified reward. The rewarded sample serves merely to indicate the choice-object whose position is correct. The selection of the choice-object is rewarded implicitly, not actually, on the first presentation. Furthermore, short non-spatial delayed reactions were solved in the 50 percent of the trials in Test IV in which only an 'expected' reward existed.

These data do not contradict Finan's 'minimum assumption' that ". . . a reënforcing agency, normally operating in the intact animal, is *reduced* by frontal ablation to a level of functioning that renders fixation more difficult." They do not, however, support the idea that the prefrontal areas have any unique role as an implicit reinforcing agency.

Difficulty in rapidly reversing symbolic associations already formed may well account in part for the degree of loss suffered by prefrontal monkeys on delayed reaction problems. We have already emphasized that part of the difficulty occasioned by the delayed reaction problem lies in the fact that the correct and incorrect stimuli frequently reverse in a single daily trial series. Harlow (**4**) has demonstrated that normal monkeys have greater difficulty in learning a series of (delayed) discriminations in which the correct and incorrect stimulus-objects frequently reverse than in learning a series comprised of new, unrepeated pairs of stimulus-objects. This difficulty encountered in normal monkeys in effecting frequent reversals to the correctness or incorrectness of a stimulus may be exaggerated by frontal lobe lesions.

Observations on a baboon (Jr, in paper I of this series) lend some support to such a hypothesis. This animal failed to solve delayed reactions longer than 5 sec. in length when a 25-trial series was presented daily. The same subject made a score of over 90 percent correct responses on a series of 25 delayed responses of 15 and 30 sec. duration when a single trial a day was run.

The findings of the present paper in conjunction with those of

related papers do not support the hypothesis that bilateral prefrontal lobectomy in the monkey destroys some single ability. Indeed, there is no conclusive evidence that any particular ability is lost and there is considerable evidence that more than one function is considerably impaired. Decrease in ability to solve delayed reactions may result from loss in speed of learning, loss in effectiveness of 'expected' reward, and difficulty in reversing symbolic associations. Other losses may be indicated by subsequent experiments.

VII. Summary and Conclusion

(1) Two rhesus monkeys, following bilateral removal of the prefrontal areas, were tested on spatial and non-spatial delayed reaction problems. These were run using a matching-from-sample technique.

(2) Both animals solved spatial delayed reactions using the matching-from-sample method of stimulus presentation. The maximal delay intervals successively bridged were 30 sec. and 10 sec. for Subjects 54 and 55 respectively.

(3) Both monkeys responded effectively to short non-spatial delayed reactions when a variation of the matching-from-sample method was used. Subject 54 solved 10-sec. non-spatial delayed reactions and subject 55 solved 5-sec. non-spatial delayed reactions.

(4) The data of this paper indicate that monkeys following bilateral removal of the prefrontal areas can respond successfully to delayed reaction problems when tested by means of a matching-from-sample technique.

(5) The data of this paper in conjunction with the data of similar researches do not support the hypothesis that bilateral prefrontal lobectomy destroys any particular ability, but suggest that a number of functions are impaired.

(Manuscript received October 22, 1942)

References

(1) Finan, J. L. Effects of frontal lobe lesions on temporally organized behavior in monkeys. *J. Neurophysiol.*, 1939, **2**, 208–226.

(2) Finan, J. L. Delayed response with pre-delay reënforcement in monkeys after the removal of the frontal lobes. *Amer. J. Psychol.*, 1942, **55**, 202–214.

(3) Harlow, H. F. Responses by rhesus monkeys to stimuli having multiple sign-values. In McNemar, Q., & Merrill, M. A. (edit.). *Studies in personality.* McGraw-Hill Book Co., 1942, 105–123.

(4) Harlow, H. F. Studies in discrimination learning by monkeys. II. Discrimination learning without primary reinforcement. *J. gen. Psychol.* (In press.)

(5) Harlow, H. F., & Dagnon, J. Problem solution by monkeys following bilateral removal of the prefrontal areas. I. The discrimination and discrimination reversal problems. *J. exp. Psychol.* (In press.)

(6) Jacobsen, C. F. A study of cerebral function in learning. The frontal lobes. *J. comp. Neurol.*, 1931, **52**, 271–340.

(7) JACOBSEN, C. F. Functions of the frontal association area in primates. *Arch. Neurol. Psychiat.*, 1935, **33**, 558–569.
(8) JACOBSEN, C. F. Studies of cerebral functions in primates. Comp. Psychol. Monogr., 1936, **13**, No. 3, pp. 68.
(9) KRECHEVSKY, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, **45**, 107–133.
(10) MALMO, R. B. Interference factors in delayed response in monkeys after removal of the frontal lobes. *J. Neurophysiol.*, 1942, **5**, 295–308.
(11) NISSEN, H. W., RIESSEN, A. H., & NOWLIS, V. Delayed response and discrimination learning by chimpanzees. *J. comp. Psychol.*, 1938, **26**, 361–385.
(12) WEINSTEIN, B. Matching-from-sample by rhesus monkeys and by children. *J. comp. Psychol.*, 1941, **31**, 195–213.

THE PRODUCTION OF AUDIOGENIC SEIZURES BY INTERRUPTED TONES *

BY ROBERT GALAMBOS AND CLIFFORD T. MORGAN

Harvard University

Introduction

In order to contribute to the understanding of the mechanism of audiogenic seizures in rats, we have been attempting to define the essential characteristics of the auditory stimuli which elicit these seizures. Auditory stimuli have three essential aspects whose relation to seizures in rats can be studied: frequency, intensity and duration of presentation of sound. In a previous paper (1), we studied the part played by frequency and intensity, concluding that, irrespective of frequency, sounds of approximately 100 db. sensation-level for the rat incite seizures. In the present study the frequency and intensity of stimulation were kept constant for all tests, and the duration of stimulation was made variable.

While there are a large number of ways in which the duration of stimulation could have been varied, all of them can be reduced to varying either (a) the *ratio* of sound to silence or (b) the length of the *cycle* of sound and silence. We chose therefore a few representative ratios and cycles and determined for them (a) the percentage of exposures in which each ratio and cycle caused seizures, and (b) the time elapsing from the beginning of exposure to the onset of the seizure, that is, the latency of the seizure.

Thus, some of the questions which this research attempted to answer are as follows: Must an auditory stimulus be maintained continuously for seizures to occur, or are interrupted tones also effective? If interrupted tones are adequate, what proportion of the time must the tone remain on, and how long may be the intervals of silence? Moreover, how does presenting an interrupted tone affect the latency of the seizures that are obtained? We believe we have been able to answer these questions in quantitative, and relatively exact, terms.

Materials and Procedure

The equipment used for generating and measuring pure tones in this study is the same as that described in a previous report by Morgan and Galambos (1), except for a device with which

* From the Department of Physiology, Harvard Medical School, and the Department of Psychology, Harvard University.

to interrupt tones. In the first of the four experiments reported below, this was accomplished by wiring a metronome with mercury contacts into the sound-producing circuit. In the other experiments, the interruption was secured by revolving a sectored disc in the path of light to a photoelectric cell controlling the sound-producing circuit. With either arrangement it was possible to interrupt tones in any convenient ratio of sound to silence.

The procedures followed in the four experiments of this study are summarized in Table 1.

TABLE 1

Experiment	Procedure						Results		
	No. of Animals	Ratio Sound to Silence	Length of Cycle (secs.)	Time On (secs.)	Time Off (secs.)	No. of Exposures	No. of Seizures	Percent Seizures	Mean Latency
I	28	Steady tone	—	—	—	140	31	22	47.5
		1 : 1	.25	1/8	1/8	140	30	21	45.5
		1 : 3	.50	1/8	3/8	140	8	6	48.7
II	64	1 : 1	.11	1/18	1/18	96	22	23	50.3
		1 : 1	.22	1/9	1/9	96	21	22	46.6
		1 : 1	.67	1/3	1/3	96	23	24	53.1
III	64	1 : 1	.67	1/3	1/3	288	50	18	49.0
		1 : 1	2.00	1	1	288	36	12	56.3
		1 : 1	10.00	5	5	288	30	10	57.9
IV	64	1 : 1	.25	1/8	1/8	96	24	25	50.8
		1 : 3	1.00	1/4	3/4	96	1	1	—
		1 : 3	.25	1/16	3/16	96	10	10	63.0

There it can be seen that sound-to-silence ratios of 1 : 0 (steady tone), 1 : 1 and of 1 : 3 were used. Several lengths of the sound-silence cycle, irrespective of ratio, were also investigated. The range covered by the 1 : 1 ratio was most extensive, the shortest cycle being 1/9 sec. (1/18 on, 1/18 off) and the longest being 10 sec. (5 on, 5 off). The range of observations with the 1 : 3 ratio, on the other hand, was from a cycle of 1/4 sec. (1/16 on, 3/16 off) to a cycle of 1 sec. (1/4 on, 3/4 off).

The standard tone used throughout all of the experiments was 4 kc. at an intensity-level of 134 ± 8 db. above 10^{-16} watts per cm.^{2} as measured with a sound-level meter in the cage where the rats were stimulated. This tone has previously been found to be serviceable for eliciting seizures (Morgan and Galambos, 1).

The rats were four to twelve weeks old at the beginning of the experiments; they had been bred from 'epileptoid' parents and were carefully selected for high susceptibility to seizures. Between experiments, those animals whose performance was poor in the experiments just run were replaced with new animals whose susceptibility to seizures had been established in preliminary tests. Twenty-eight rats took part in the first experiment and 64 in subsequent experiments, but because of occasional replacements between experiments the total number of animals used was about 100.

The rats were exposed on one day each week in groups of four. This unit of four rats was counted as four exposures. To arrive at the 140 exposures tabulated in Experiment I, for example, seven groups of four rats were stimulated five times with each of the stimulus-conditions indicated. In later experiments two main groups of 32 animals each and of about equal susceptibility to seizures were employed, but each of these main groups did not necessarily receive the same number of exposures. Thus, in Experiment II, one group of 32 animals was stimulated in units of four only once with each stimulus-condition, while a second group of 32 was stimulated in such units of four two times with each condition, making 96 exposures in all.

Each group of four rats was presented with three stimulus-conditions, as indicated in Table 1, on each day of testing. To guard against adaptation entering into our results undesirably, the

order of presentation was varied in such a way that each stimulus-condition occurred first, second and third an equal number of times. Each exposure was at least 90 sec. in length, and 20 min. was allowed to elapse between the end of one exposure and the beginning of the next. If an animal had an attack in one exposure, it was nevertheless stimulated again in the stimulus-condition which followed, provided its attack did not occur on the third exposure of a day. It is probable that having an attack in one test affects susceptibility in the next, but since the order of presentation of different stimulus-conditions was counterbalanced to neutralize such factors as adaptation or changing suceptibility, this factor should not affect our results in any systematic way.

There was little difficulty in deciding whether or not a seizure occurred, and when it occurred. The criterion was always that there be a period of frenzied, violent running, lasting at least 5 sec. Premonitory symptoms, such as face-washing, weaving back and forth, and short running periods, and startle responses upon the initial presentation of the sound were not counted. Also, only those seizures which occurred within the first 90 sec. after the beginning of an exposure were considered.

Results

Experiment I.—Table 1 gives a summary of all the relevant data in this experiment. The three stimulus-conditions were: a steady tone; 1/8 sec. on and 1/8 sec. off; and 1/8 sec. on and 3/8 sec. off. A total of 140 exposures were run with each stimulus-condition. About the same percentage of seizures occurred with the 1/8 : 1/8 tone as with the steady tone; however, significantly fewer seizures were obtained with the 1/8 : 3/8 condition. The latencies of attacks occurring under the different conditions, on the other hand, were the same in all cases—a very noteworthy fact in view of the different amounts of total sound energy delivered in the three stimulus-conditions.

Two aspects of the interrupted tones of this experiment might be responsible for the differences in their effectiveness: (a) the ratios of sound to silence, which in one was 1 : 1 and in the other 1 : 3, and (b) the length of the silent interval, which in one was 1/8 sec. and in the other 3/8 sec. The subsequent experiments were designed to determine what part these two factors play in the difference in effectiveness of interrupted tones.

Experiment II.—In this and the following experiment, the ratio of sound to silence was kept constant at 1 : 1 and the duration of the cycle (sum of sound and silent intervals) was varied from very short to relatively long values. In Experiment II, the cycle-lengths were 1/9, 2/9 and 2/3 sec. As is shown in Table 1, 96 exposures were given at each of these cycle-lengths, and almost precisely the same percentages of seizures were obtained. Thus, the duration of the silent interval, within the limits specified, is not significantly related to the number of seizures obtainable. Because the 1/9 : 1/9 stimulus is approximately the same as the 1/8 : 1/8 stimulus which was shown in Experiment I to be as effective as the steady tone, we may conclude that all the interrupted tones of 1 : 1 ratio and of

2/3 sec. sound-silence cycle or shorter are approximately as effective as a steady tone. It will be noted, finally, that the latencies of attacks given in Table 1 for the three different stimulus-conditions of Experiment II are not significantly different from each other.

Experiment III.—In this experiment the cycle durations of 2/3, 2.0, and 10.0 sec. were presented in the ratio of 1 : 1. This experiment, therefore, is an extension of Experiment II to longer cycle duration. From Table 1 it seems clear both that longer cycle durations constitute less effective stimuli, and that a progressive though slight increase in the latent period occurs with increase in cycle duration. It will be noted that a larger number of exposures (288) were given in this experiment than in any other. This was necessary in order to cancel the effects of the somewhat wider day-to-day variations in the experimental results than were evident in the other experiments.

Experiment IV.—In this experiment, summarized like the others in Table 1, the effectiveness of the 1 : 3 ratio at cycle durations of 1.0 and 1/4 sec. was compared to the 1 : 1 ratio at a cycle of 1/4 sec. The results obtained show that neither 1 : 3 ratio is as effective as the 1 : 1; furthermore, the longer cycle durations appear to be less effective in the 1 : 3 ratio just as in the 1 : 1 ratio (see Exp. III).

Discussion

As was previously indicated, care was taken in each of the four experiments to counterbalance the order of presentation of different stimuli so that any systematic changes from one test to another in the day, such for example as adaptation of animals to the effects of stimulation or to the effects of having had attacks, would not adversely affect the results. No such arrangement, however, was followed with respect to the comparison of different experiments with each other. Indeed, animals that were not giving seizures in one experiment were dropped out and others added before beginning the following experiment. These considerations would not permit us to make comparisons between different experiments.

It happens, however, that there was always one stimulus-condition which was common to two or more of the experiments. Thus, 1/8 sec. on, 1/8 sec. off, appears in Experiments I and IV eliciting 21 and 25 percent seizures respectively; its approximate equal, 1/9 sec. on, 1/9 sec. off, was employed in Experiment II with 22 percent seizures. Finally, 1/3 sec. on, 1/3 sec. off, appears in Experiments II and III with 24 and 18 percent seizures respectively.

These differences in the percentage of attacks obtained with the same stimulus-conditions are not large and are within the range of

variation that might be expected if, despite changes in the population between experiments, susceptibility were the same in the different experiments. We have, accordingly, considered it legitimate to assume that the animals were of comparable susceptibility in the different experiments and have made comparisons between the experiments. Thus, in Fig. 1, all of the results obtained for the 1 : 1

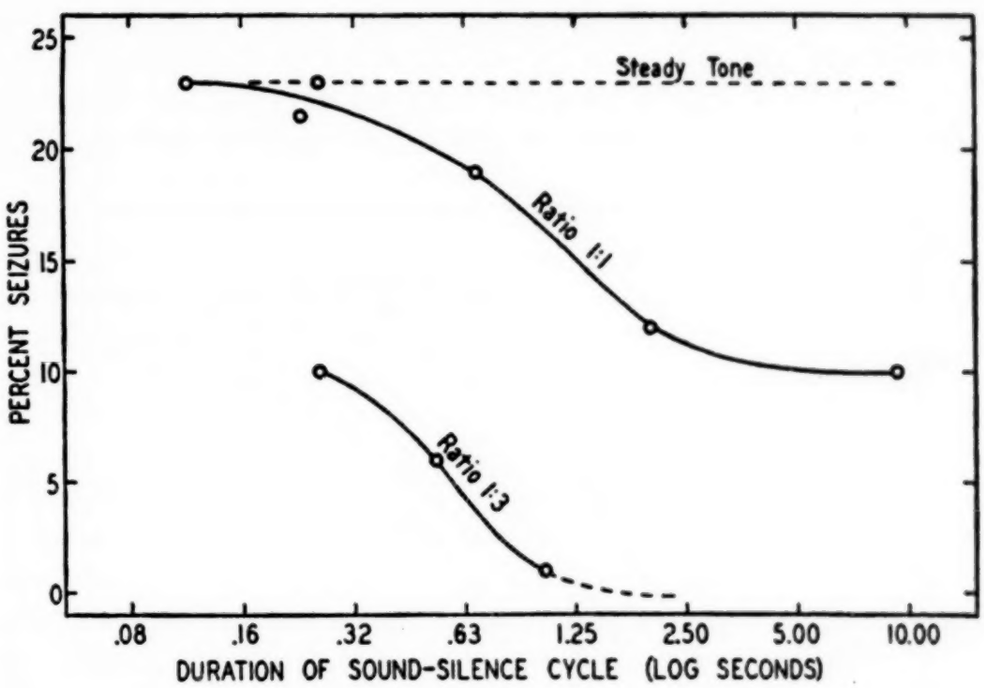

FIG. 1. Effectiveness of different interruptions of a 4 kc. tone at 134 ± 8 db. in eliciting audiogenic seizures in rats. This plot summarizes the data in Table 1. The ratio of sound to silence in the interrupted tone is the parameter.

ratio of sound-to-silence have been plotted on one curve and all those for the 1 : 3 ratio have been fitted with another curve.

We may consider first the relationship of the cycle duration (sum of sound and silent intervals) to the incidence of seizures, holding the ratio as parameter. Fig. 1 shows the results of the four experiments performed plotted so as to show the percentage of seizures occurring at each sound-silence cycle at ratios of 1 : 1 and 1 : 3. In cases where two determinations for a particular cycle duration were available from different experiments, these were averaged (taking care to consider the number of exposures upon which they were based) and plotted as one point (see values of .25 and .67 for the 1 : 1 ratio). Free-hand curves were then fitted to the resulting distribution of points.

It can be seen that the resulting curve is relatively smooth and gives a clear picture of the relationship between cycle duration and percentage seizures. As the length of the cycle is increased, the

percentage of seizures falls for both the 1 : 1 and the 1 : 3 ratios. In the 1 : 1 curve, the relationship appears to be sigmoid, when percentage is plotted against log time; it falls from a maximum of 23 percent at cycle-lengths below .20 sec. to 10 percent at cycle-lengths above 5.00 sec. In the case of the 1 : 3 ratio, it is not clear whether the curve is sigmoid in shape or not, for determinations at the shorter cycle lengths are lacking. It is clear, however, that effectiveness is reduced to zero when the cycle is about 2.0 sec. or longer in duration.

It is clear from Fig. 1 that the ratio of 1 : 3 (sound to silence) is uniformly less effective than is the 1 : 1 ratio at any given cycle

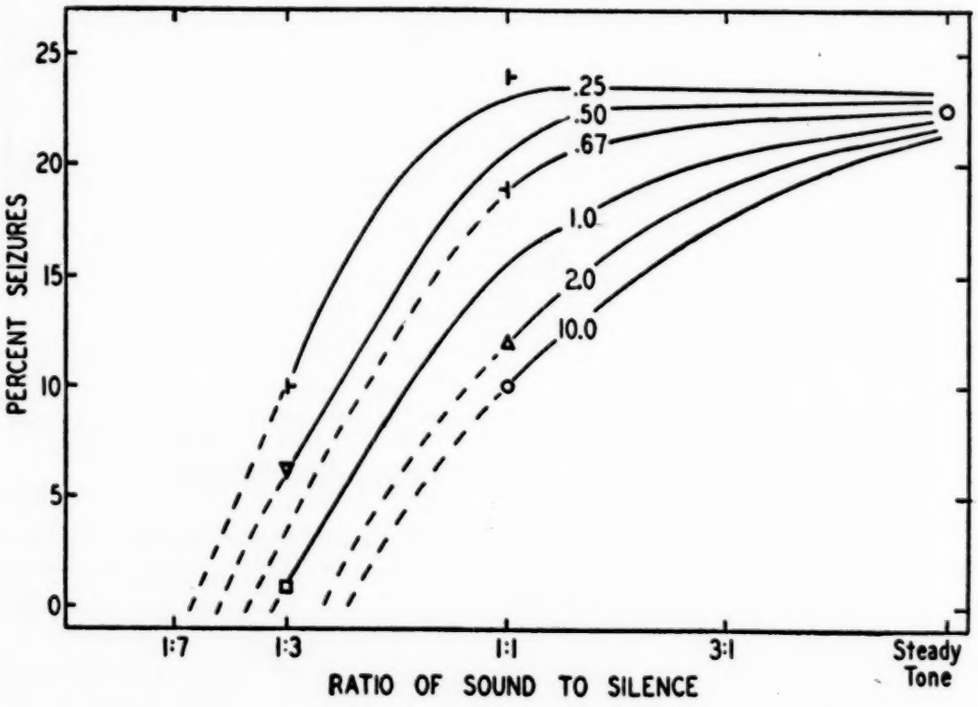

FIG. 2. Effectiveness of different interruptions of a 4 kc. tone at 134 ± 8 db. in eliciting audiogenic seizures in rats. In this figure the data in Fig. 1 are replotted so as to make cycle-length (sum of sound and silence intervals) the parameter.

duration. A more comprehensive picture, however, can be gotten by replotting the curves of Fig. 1 so as to make cycle duration the parameter and ratio the variable. This has been done in Fig. 2, where the proportion of a cycle during which the sound is on has been graphed with respect to percentage of seizures it elicits. There it is shown that for cycles of about .20 sec. (*e.g.*, 1/9 on, 1/9 off) sound-to-silence ratios of 1 : 1 or greater are approximately as effective as a steady tone, and there is no effect of interruption upon the incidence of seizures. With smaller ratios, however, the effectiveness is rapidly reduced, and if our extrapolations of the curves are substantially correct, no seizures at all should be expected at ratios

of 1 : 7 or less when the cycle is .25 sec. That is to say, an interrupted sound of 1/8 on and 7/8 off should be ineffective no matter how rapidly the tone is interrupted. For the 10-second cycle, as opposed to very short ones, the percentage of seizures decreases very rapidly as the ratio of sound to silence is decreased; at a ratio of 1 : 1 fewer than half as many seizures are obtainable as with a steady tone. Moreover, by extrapolation, it may be predicted that a ratio of 1 : 3 or less will be completely ineffective for these long cycles. Curves for cycles between the extremes of .25 and 10 sec. just cited fall at intermediate positions as shown in Fig. 2.

The significance of these findings falls, we believe, into the following categories. First, we believe it is now possible to predict the effectiveness of any tone of 4 kc. at 134 ± 8 db. in our colony of rats. This follows from Fig. 2, where any point within the area enclosed by the solid lines not only designates a particular stimulus-condition, defined in terms of cycle duration and ratio of sound to silence, which causes seizures, but also it tells how effective that stimulus will be in terms of percentage of seizures obtained. To be sure, Fig. 2 refers only to 4 kc. tones at 134 ± 8 db.; but similar relationships could undoubtedly be demonstrated at other frequencies and intensities of stimulating tones.

The second fact which these data demonstrate is the relative independence of the response of the animal to the amount of stimulus energy delivered. Twice as much energy is delivered by a steady tone as by one consisting of 1/9 sec. sound and 1/9 sec. silence. Yet the steady tone is no more effective than the interrupted, either in terms of causing more seizures or in evoking them sooner.

The relative constancy of the latent period under all the varied conditions of stimulation implies that those events which determine audiogenic seizures require a relatively fixed amount of time for their completion. Apparently any adequate stimulus sets off changes within the animal which culminate in the seizure; increase in intensity of stimulation cannot significantly modify the fixed pace at which these changes occur. There seems to be a trend toward longer latency with longer cycle length (see Table 1), but the differences are slight and statistically insignificant.

Summary

The effectiveness of intermittent sound stimulation has been studied in an attempt to define further the stimulus conditions required for eliciting audiogenic seizures in rats. The stimulus used in the present experiment was a 4000-cycle tone at 134 ± 8 db. (above 10^{-16} watts per cm.^{2}), interrupted in such a way that a known

interval of silence followed a known interval of sound. The duration of the sound-silence cycle was varied as well as the ratio of sound to silence.

At a sound-to-silence ratio of 1 : 1, the stimulus becomes more effective as the sound-silence cycle decreases from 10 sec. (5 sec. on, 5 sec. off) to about 1 sec. (1/2 sec. on, 1/2 sec. off). Cycles shorter than 1 sec. are just as effective as a steady tone in eliciting seizures.

At a ratio of 1 : 3, the 1-second cycle (1/4 sec. on, 3/4 sec. off) is practically ineffective, but cycles shorter than this cause progressively more seizures. At a given cycle duration the 1 : 1 ratio always elicits more seizures than the 1 : 3 ratio.

The latent period for the seizure was approximately the same for all the different types of stimuli.

(Manuscript received January 14, 1943)

References

1. Morgan, C. T., & Galambos, R. The production of audiogenic seizures by tones of low frequency. *Amer. J. Psychol.*, 1942, **60**, 555–559.

AN EXPLORATORY STUDY OF PUPILLARY RESPONSES DURING DECEPTION

BY F. K. BERRIEN AND G. H. HUNTINGTON

Colgate University

It has long been known that pupillary changes accompany emotional disturbances. As early as seventy-five years ago the pupil was considered by at least one physiologist, Schiff, to be the finest esthesiometer of the body. Furthermore, our contemporary medical and psychological clinicians usually have been mindful of the diagnostic significance of pupillary symptoms in a large variety of illnesses, especially those with affective complications.

In spite of this long established and wide-spread recognition of the importance of pupillary responses in emotions, it has only been within recent years that precise systematic studies of these responses have been made. For the most part, our knowledge of pupillary responses, particularly during sudden fright or startle, has been based upon direct observation under the varied conditions of a clinical examination. Partly because of the crudeness of these observations and the lack of permanent records, certain misconceptions concerning pupillary responses were to be expected. However, a number of new techniques have been developed in the past decade, the most elaborate probably being the photographic recording equipment described by Lowenstein and Friedman (1). These and similar devices have made it possible to control the experimental situation to a high degree and at the same time obtain highly accurate records of the second-by-second action of the pupil.

The present writers have been interested in discovering, first, what kinds of pupillary responses accompany the tension and emotional excitement incident to deception and second, whether these responses show sufficient consistency in any respect to permit them to be used as valid indications of attempted deceit. Furthermore, since considerable validating evidence is already at hand both from laboratory and field studies pertaining to the use of blood pressure changes as indicators of deception, a comparison of these changes with pupillary responses was proposed.

Apparatus

The apparatus consisted of a constant speed polygraph using capillary pens for purposes of recording: (a) the moment when questions were given and answers received; (b) the pulse wave;

and (c) the pupillary response. The pulse record was obtained by using a standard blood pressure cuff wrapped around the left arm and inflated to a pressure sufficient to give a clear graphic record, the pressure usually being 70 to 80 mm. This cuff was connected to a balanced pressure metal recording bellows whose recording lever activated a light-weight capillary pen. The pupillary responses were recorded by means of a short-focus telescope mounted on a horizontal rack and pinion in such a manner that the vertical cross hair could be brought tangent to the *S*'s pupil as seen in the field of the telescope. As the pupil dilated or constricted, the telescope was moved right or left by the observer; the action of the telescope was transmitted to, and amplified mechanically by, a recording pen. It was, of course, necessary to hold the *S*'s head rigid by means of a comfortable head clamp. The *S*'s fixation point was 22 in. from his eye. A 10-watt incandescent lamp enclosed in a dome-shaped metal shade covered with architect's cloth was located beside the *S*'s face so that the pupil was adequately illuminated and yet moderately dilated. A pnemograph was strapped around the *S*'s chest but the records of breathing are not considered in this report.

Subjects and Procedure

The *S*s consisted of more than 40 college students. Only 32 of the records were satisfactory for purposes of this research, mechanical failures or irregularities in following instructions making it necessary to discard the others. The *S*s reported to the laboratory in pairs and were told that the experiment was one in lie detection. They were asked to follow further instructions which could be found in envelopes at a designated place. Each *S* selected one of the two sets of directions out of the experimenter's sight. One set of instructions told the recipient that he was innocent and should merely go into the hallway and wait; he should not follow the other *S*, and upon questioning he should tell the truth on all items. The other *S* was instructed to go to a designated spot in a lecture hall, to take some money left there (seldom more than a dime), to return to the laboratory without telling his partner what he had done, and to lie on all questions connecting him in any way with the 'crime.' Moreover, if he were able to deceive the experimenter, he could keep the money. Both men returned to the laboratory together but were taken singly into the experimental room.

While one experimenter adjusted the apparatus to the *S*, the instructions for the examination itself were given. The *S* was told there would be two series of 6 or 7 questions each with a rest pause between. The first was merely a series of control questions having nothing to do with the crime and designed only to find his 'normal' reactions to the apparatus. He was told that if he were guilty, he would have nothing to worry about on this first series. The second series of questions however, would deal with the crime. He was further instructed to delay his answers to all questions until one of the experimenters said, "Now."

In order to increase the emotional effect of lying, the critical questions pertaining to the crime were actually introduced after only two or three control questions. Some preliminary experiments had demonstrated that the critical questions were interpreted by the innocent individuals as merely part of the regular controls, while the guilty were caught somewhat off guard and consequently had their emotional tension increased by reason of the surprise element.

Earlier experiments had revealed that the fixation necessary for making accurate records of pupillary changes could not be prolonged without ocular discomfort. To avoid complicating the situation by possible reactions against the experimental procedures, the questioning was limited to only three critical questions: "Were you in the lecture hall a moment ago? Did you take the money? Have you been telling me the truth?"

The delay in answering was imposed upon the *S*s in order to provide time during which the tension stemming from the question itself could be recorded. This arrangement tended to prolong any tension so that it could be clearly recorded. The signal for answering was given after three complete respiratory cycles on the part of the *S*.

Results

The most obvious way in which to evaluate the validity of the combined pupillary and blood pressure records is to ask, "How many correct diagnoses of guilt and innocence were made based on a study

of the records alone?" After each pair of *S*s had been questioned, the senior experimenter examined the records and attempted to identify the 'criminal.' Only those cases were included which failed to 'give themselves away' by their verbal responses, or in other obvious ways. The senior experimenter was further prohibited from gaining other cues as to guilt or innocence by reason of his close observation of the *S*'s pupil throughout the questioning. In the 16 pairs of records thus examined three mistakes were made identifying the 'criminal.' These diagnoses, of course, represented an over-all interpretation of the records, a task comparable to that thrust upon the criminal investigator, except that the cues here were limited to the records alone, uncomplicated by circumstantial evidence and other contributing factors. It must be emphasized also that some skill in interpreting the pupillary records had been developed through the previous two years when the present apparatus and procedures were being developed.

In attempting to analyze the particular cues employed in making these diagnoses, one is impressed by the extreme variability of all records and the apparent absence of any clear consistent pattern of response either among those telling the truth or those attempting deceit. With regard to the pupillary response, approximately 50 percent of the cases showed a gradual 'drift' from a moderately dilated pupil to a more constricted state as the examination progressed, but this was no more characteristic of the guilty than the innocent. It was apparently a reflection of a gradual adaptation of the subjects to the experimental situation.

Upon careful examination, a unique pattern of response was discovered in a large number of cases in which deceit was attempted. This pattern, as illustrated in Fig. 1 at point A, appeared in only a small number of truthful responses. This response pattern can be described as possessing two phases, the first being a slow, and usually steady dilation negatively accelerated. This phase may last for approximately one to five seconds, depending partly upon the magnitude of the change. In the second phase, the pupil constricts very rapidly to a point approximately equal to its original size, although under continued emotional tension it frequently reverses its action before reaching that point. Such a response pattern occasionally appears 'spontaneously,' *i.e.*, between questions, much as one finds stray deflections in galvanic recordings of the electro-dermal phenomena. It also appears in control questions, especially following those questions coming early in the examination, and it may therefore be considered evidence of the *S*'s initial apprehension and lack of adjustment to the experimental situation.

The accompanying table indicates the frequency and distribution

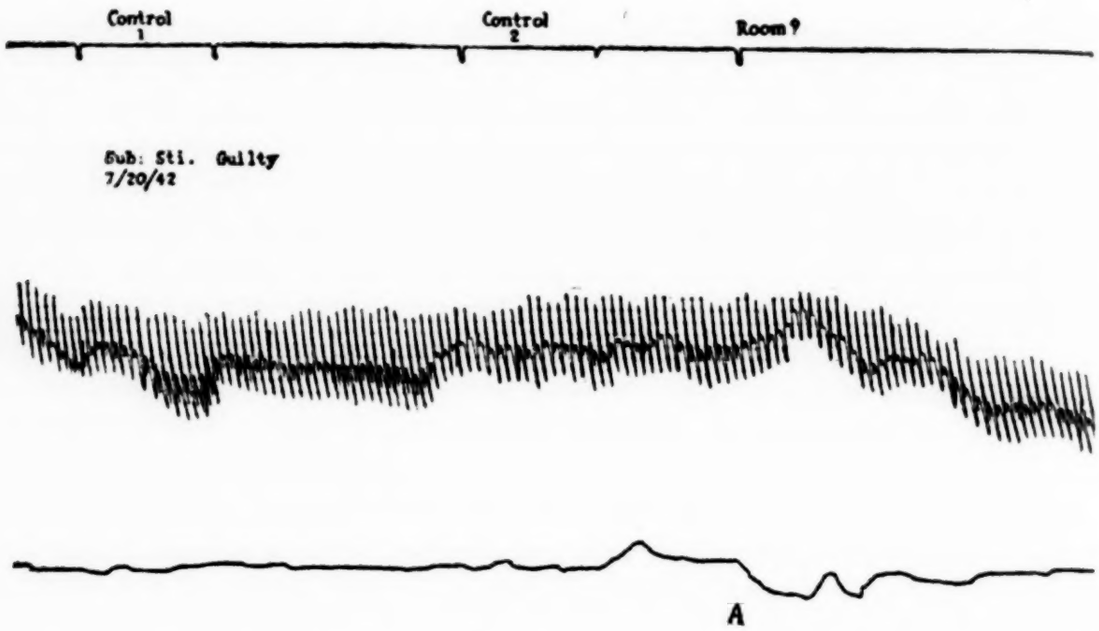

FIG. 1. Illustration of pupillary response during attempted deception, showing slow and negatively accelerated dilation (downward on the record) and rapid constriction (upward on the record).

of this type of response occurring either immediately after the question or after the *S*'s response. In many but not all of the cases indicated in the table, the unique pupillary response occurs after both the question and the answer.

A second feature of somewhat less importance in distinguishing between 'criminal' and 'non-criminal' records is the change in the stability of the pupil. In the cases indicated by the asterisk, the pupil remained quite constant in size until the crucial questions were introduced and then suddenly began to vary in size over a much wider range than previously. As the table indicates, this appears to occur more frequently among the guilty although, like the previously described response, it is not an infallible symptom.

Considering the guilty *S*s in the table, it is evident that two failed to give pupillary symptoms of unusual emotional tension. Possibly the experimental set-up was not sufficiently exciting to elicit emotional responses in these *S*s. On the other hand, the blood-pressure records of one of these *S*s showed a typical deceit response pattern, but the other did not. Among the innocent at least four (Tin., Smi., Car., and Mil.) would have been considered likely 'criminals' had the pupillary records been the sole basis for judgment.

A similar analysis of the blood pressure records was made in which significant changes were tabulated for each *S*'s response to each question. This tabulation revealed increases in pressure on 36 of the 64 control questions asked of the entire group, increases being distributed almost equally among the innocent as well as the guilty.

This compares with only 11 pupillary responses indicative of emotion on these questions.

On the basis of this finding one suspects that the vascular system is somewhat more sensitive to all emotional stimuli, although the slow-dilation-quick-constriction of the pupil appears to be somewhat more limited to the kind of emotion that accompanies deceit.

Interpretation of the blood pressure records independently of the pupillary responses led to a misidentification of the 'criminal' in four cases (Bri., Car., Mil., and War.), the distinguishing symptom being a greater-than-normal rise in the base line of the pulse wave on critical questions. The normal amount of rise was judged for each *S* on the basis of his control questions. Only two of the latter misidentifications (Car. and Mil.,) are the same as those made on the basis of the pupillary responses. These two *S*s were also among the three mistakenly identified as 'criminals' at the time of the recordings. Bri. was the third incorrectly identified 'criminal.' Since all of the three misidentifications at the time of the examination had pulse records suggestive of deceit, it is evident that these over-all diagnoses were more influenced by the blood pressure than the pupillary indications. Had the reverse been true probably the same number of mistakes would have occurred but with different *S*s.

Inspection of the records shows clearly that there is little correspondence between the action of the pupil and vascular pressure

TABLE I

OCCURRENCE OF PUPILLARY RESPONSES INDICATIVE OF EMOTION

Guilty Subjects	Control Questions		Critical Questions			Innocent Subjects	Control Questions		Critical Questions		
	1st	2nd	1st	2nd	3rd		1st	2nd	1st	2nd	3rd
For.			x		x	Pla.					x
Ban.			*	*		Mul.					
Han.				x	x	Tin.				x	x
Thu.						Pet.	x	x			
Vol			x	*	x	Bri.	x		x		
Tho.			*	*	x	Smi.	x	x		x	x
Fie.	x	x	x	x	x	And.					x
Col.	x		x			Car.	x		x	x	
Gra.						Mil.		x	x	x	x
Far.			x	x	x	O'De					
Sti.			x			War					
Tho.			x	x	x	Pet.	x		x		
Bus.			*	*		Rei.					
All.				x		DeB.					
Fly.			x	x	x	Dec.			*	*	*
O'B.				x	x	Hec.					

Slow-dilation—quick-constriction of pupil is indicated by x.
Markedly increased variability of pupil compared with control questions is indicated by *.

changes as we recorded them. In other words, the pupil rarely dilates or constricts in step with blood pressure changes, all visual stimulation remaining constant. Moreover, of course, the pupil responds much faster than the vascular system to ideational stimuli. This lack of correspondence is borne out in part by the fact that misidentifications of the 'criminals' by means of the pupillary responses were not the same as those made on the basis of the blood pressure changes. At the same time, the 'symptoms of deceit' in the blood pressure and the pupillary records occur together in 48 percent of the responses when deceit is attempted and in only 18 percent when deceit is not attempted. Consequently, while there is little correspondence between the records, one complements the other in revealing emotional tension with sufficient frequency to make them both useful in the diagnosis of deceit.

Discussion

The procedure used in this study was necessarily rigid for purposes of research. Had the questioning been varied to fit each *S* by means of a greater number of questions, difficult cases might have revealed their guilt or innocence more clearly. However, such a procedure would have been as much a test of the experimenter's skill in examining as a test of the apparatus.

Support for the hypothesis that the slow-dilation—quick-constriction pattern is an emotional response is not confined to this study alone. Löwenstein and Friedman (1) found a type of response very similar, but of a shorter duration, when patients suffering from hysteria with somatic symptoms were stimulated by a brief flash of light. Even though in the present study we are dealing with ideational rather than physical stimuli and with transitory rather than chronic emotions, the similarity in response patterns is remarkable.

Greater success in identifying the 'criminals' was probably limited by a number of factors. First, many of the *S*s assumed a 'game attitude' making them more amused than tensely excited. Second, in spite of the effort to catch the criminal off guard, there was little to distinguish the reported feelings of innocent *S*s from the guilty. This is probably inevitable in any controlled laboratory arrangement capable of repetition. A greater monetary lure might have been used to increase the emotional tension. The percentage of correct identifications was somewhat lowered because we made diagnoses in cases which were not clear-cut. No cases of indecision were permitted. This would not be the case in field applications of similar techniques, where at present it is common to withhold a diagnosis of either guilt or innocence in roughly 20 percent of the cases.

The fact that trickery was employed in this experiment to enhance

the emotional tension of the 'criminals' should not be construed as a recommendation for such a procedure in bona-fide criminal investigations. The ruse was employed solely as a laboratory devise to create a degree of emotion somewhat comparable to that present in persons under genuine suspicion with more than a dime at stake.

Summary and Conclusions

The foregoing study was designed to discover whether any consistent pattern of pupillary response is associated with deception. Incidental to that purpose, the study also attempted to compare the diagnostic value of blood-pressure changes and pupillary responses as indicators of deception. On the basis of the 32 records considered in this paper it is evident that:

(1) A slow dilation followed by a very rapid constriction is probably indicative of the emotion usually accompanying deceit.

(2) A sudden change in the stability of the pupil is found more frequently among those attempting deceit than those not attempting deceit when critical questions are first introduced.

(3) The pupillary records are no more indicative of deceit than are the pulse records, each giving the same percentage of error in identifying the 'criminals.' When the identification is based on interpretation of the combined records, an error of approximately 20 percent is obtained.

(4) The pupillary responses do not ordinarily correspond in any way with the vascular responses recorded in this study. However, in 48 percent of the falsehoods both the vascular system and the pupil give indications of emotional tension. This compares with 18 percent when truthful answers are given to critical questions.

(5) A number of known factors operated to decrease the number of correct identifications in the laboratory. It is possible that in the application to criminal investigations the techniques described may prove slightly more valid than the present study indicates.

(Manuscript received December 17, 1942)

Reference

1. Löwenstein, O., & Friedman, E. D. Pupillographic studies. 1. Present state of pupillography; its method and diagnostic significance. *Arch. Ophthal.*, 1942, 27, 969–993.

THE CONSTRUCTION OF A TILTING ACTIVITY CAGE

BY W. A. BOUSFIELD AND F. A. MOTE, JR.

University of Connecticut

The usefulness of devices for measuring the level of activity of laboratory animals is indicated by their construction in a wide variety of designs. Kreezer (1) has recently reviewed the types of apparatus employed in the measurement of the activity of the rat and has classified these types as follows: (1) vertically revolving drums; (2) horizontally rotating turntables; and (3) tambour or spring mounted activity cages. We describe below a tilting type cage which does not resemble the commonly used devices. Some of its advantages are ease and simplicity of construction and inexpensiveness; one can be built in a few hours and the cost per instrument should not exceed $3.00. A battery of these instruments has been in use in our laboratory for several months and the records obtained on rats show reasonable reliability.

As shown in Fig. 1 the cage itself is a rectangular wire-mesh rat trap, 18 in. x 7 in. x 8 in., with the central partition removed and the entrance closed. The hemispherical food basket is retained and serves to hold a supply of food and a medium sized water bottle. A wooden cradle is constructed to hold the cage, and is mounted on an axis (a wood dowel) to allow tilting as the animal moves from one end to the other. Cushions attached to the base of the apparatus

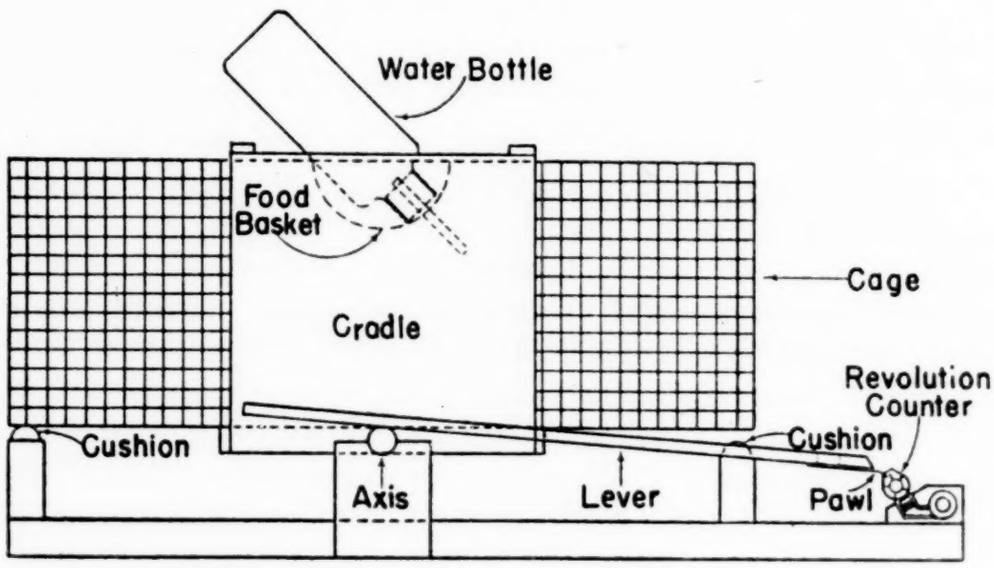

FIG. 1. A tilting activity cage.

limit the amount of tilting of the cage to the minimum necessary for operating the recording mechanism. The latter device is a revolution counter mounted as shown in the figure; it records the number of times the cage tilts in a clockwise direction. A pawl mounted on a lever engages the sproket of the revolution counter with each clockwise movement of the cradle. The pawl is a piece of light clockspring with one end bent, and is attached to the end of the lever by means of two pins. This arrangement limits the movement of the pawl as it engages the tooth of the sprocket in a downward thrust, but allows the pawl to slip by the tooth as the lever moves upward. In attaching the lever to the cradle it is necessary to make sure that the pawl is in correct alignment with the sprocket of the revolution counter. Once the adjustment has been made, and the lever fastened in place, it should be possible to remove the cradle and cage, and return them without interfering with the engagement of the pawl with the sprocket. If the alignment is disturbed, however, it is generally possible to correct it by moving the cage slightly backward or forward in the cradle, or by adjusting the position of the revolution counter. The latter correction is made by loosening the screw holding the revolution counter to its supporting block. With the apparatus in operation the base is covered with litter or several thicknesses of newspaper to catch the debris dropping from the cage. When the apparatus is being cleaned, the cradle and cage are lifted from the base.

(Manuscript received January 4, 1943)

Reference

1. Kreezer, G. L. In: Griffith, J. Q., & Farris, E. J. (eds). *The rat in laboratory investigation.* Philadelphia: Lippincott, 1942, 199–273.